

From: Swope, Sheridan  
Sent: Thursday, June 09, 2005 4:34 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/791,980

For 10/791,980, pls search and interference search:

SID 6 against the NT and AA data bases.

Sheridan Swope, Ph.D.  
Patent Examiner, AU 1652  
Recombinant Enzymes  
571-272-0943 (voice)  
E02B71 Remsen Bld (Office)  
E02C70 Remsen Bld (Mailbox)

Ben Skars

6-520 aa  
CB

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: \_\_\_\_\_ AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compen Ltd.

## OM protein - protein search, using bw model

Run on: June 13, 2005, 18:01:52 ; Search time 157 Seconds

(without alignments)  
1269.643 Million cell updates/sec

Title: US-10-791-980-6

Perfect score: 2834  
Sequence: 1 MWARVGLLRALQLLWGLH.....WATELFWGCMWANSGLAF 520Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

```
1: Published Applications_AA.*
2: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubppaa/PCTUS_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppaa/US10E_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubppaa/US10F_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubppaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 2834  | 100.0       | 520    | 10    | US-09-862-631-6   |
| 2          | 2472  | 87.2        | 520    | 9     | US-09-737-353-2   |
| 3          | 2472  | 87.2        | 520    | 9     | US-09-950-510-2   |
| 4          | 2472  | 87.2        | 520    | 9     | US-09-950-510-21  |
| 5          | 2472  | 87.2        | 520    | 9     | US-09-950-510-24  |
| 6          | 2472  | 87.2        | 520    | 14    | US-10-028-072-144 |
| 7          | 2472  | 87.2        | 520    | 14    | US-10-140-808-144 |
| 8          | 2472  | 87.2        | 520    | 14    | US-10-121-048-144 |
| 9          | 2472  | 87.2        | 520    | 14    | US-10-123-904-144 |
| 10         | 2472  | 87.2        | 520    | 14    | US-10-140-470-144 |
| 11         | 2472  | 87.2        | 520    | 14    | US-10-175-746-144 |

|    |      |      |     |    |                    |                   |
|----|------|------|-----|----|--------------------|-------------------|
| 12 | 2472 | 87.2 | 520 | 14 | US-10-176-918-144  | Sequence 144, App |
| 13 | 2472 | 87.2 | 520 | 14 | US-10-176-921-144  | Sequence 144, App |
| 14 | 2472 | 87.2 | 520 | 14 | US-10-137-865-144  | Sequence 144, App |
| 15 | 2472 | 87.2 | 520 | 14 | US-10-140-474-144  | Sequence 144, App |
| 16 | 2472 | 87.2 | 520 | 14 | US-10-142-431-144  | Sequence 144, App |
| 17 | 2472 | 87.2 | 520 | 14 | US-10-143-114-144  | Sequence 144, App |
| 18 | 2472 | 87.2 | 520 | 14 | US-10-050-216B-9   | Sequence 9, Appl1 |
| 19 | 2472 | 87.2 | 520 | 14 | US-10-142-419-144  | Sequence 144, App |
| 20 | 2472 | 87.2 | 520 | 14 | US-10-123-262-144  | Sequence 144, App |
| 21 | 2472 | 87.2 | 520 | 14 | US-10-142-443-144  | Sequence 144, App |
| 22 | 2472 | 87.2 | 520 | 14 | US-10-121-060-144  | Sequence 144, App |
| 23 | 2472 | 87.2 | 520 | 14 | US-10-141-755-144  | Sequence 144, App |
| 24 | 2472 | 87.2 | 520 | 14 | US-10-143-032-144  | Sequence 144, App |
| 25 | 2472 | 87.2 | 520 | 14 | US-10-123-108-144  | Sequence 144, App |
| 26 | 2472 | 87.2 | 520 | 14 | US-10-123-236-144  | Sequence 144, App |
| 27 | 2472 | 87.2 | 520 | 14 | US-10-123-261-144  | Sequence 144, App |
| 28 | 2472 | 87.2 | 520 | 14 | US-10-124-819-144  | Sequence 144, App |
| 29 | 2472 | 87.2 | 520 | 14 | US-10-140-921-144  | Sequence 144, App |
| 30 | 2472 | 87.2 | 520 | 14 | US-10-140-928-144  | Sequence 144, App |
| 31 | 2472 | 87.2 | 520 | 14 | US-10-121-045-144  | Sequence 144, App |
| 32 | 2472 | 87.2 | 520 | 14 | US-10-123-292-144  | Sequence 144, App |
| 33 | 2472 | 87.2 | 520 | 14 | US-10-123-903-144  | Sequence 144, App |
| 34 | 2472 | 87.2 | 520 | 14 | US-10-124-819-144  | Sequence 144, App |
| 35 | 2472 | 87.2 | 520 | 14 | US-10-124-832-144  | Sequence 144, App |
| 36 | 2472 | 87.2 | 520 | 14 | US-10-140-925-144  | Sequence 144, App |
| 37 | 2472 | 87.2 | 520 | 14 | US-10-160-498-144  | Sequence 144, App |
| 38 | 2472 | 87.2 | 520 | 14 | US-10-124-825A-144 | Sequence 144, App |
| 39 | 2472 | 87.2 | 520 | 14 | US-10-127-829A-144 | Sequence 144, App |
| 40 | 2472 | 87.2 | 520 | 14 | US-10-127-835A-144 | Sequence 144, App |
| 41 | 2472 | 87.2 | 520 | 14 | US-10-127-839A-144 | Sequence 144, App |
| 42 | 2472 | 87.2 | 520 | 14 | US-10-127-901A-144 | Sequence 144, App |
| 43 | 2472 | 87.2 | 520 | 14 | US-10-128-633A-144 | Sequence 144, App |
| 44 | 2472 | 87.2 | 520 | 14 | US-10-131-813A-144 | Sequence 144, App |
| 45 | 2472 | 87.2 | 520 | 14 | US-10-131-818A-144 | Sequence 144, App |

## ALIGNMENTS

```
RESULT 1
US-09-862-631-6
; Sequence 6, Application US/09862631
; Publication No. US20030032164A1
; GENERAL INFORMATION:
; APPLICANT: Holmgren, Erik
; APPLICANT: Kihlen, Mats
; APPLICANT: Wood, Tim
; APPLICANT: Ekblom, Jonas
; TITLE OF INVENTION: Novel Matrix Metalloproteinases
; FILE REFERENCE: 00014REGUS
; CURRENT APPLICATION NUMBER: US/09/862,631
; CURRENT FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 206119
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-631-6

Query Match      100.0%; Score 2834; DB 10; Length 520;
Best local Similarity 100.0%; Pred. No. 2.3e-234;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MWARVGLLRALQLLWGLHDAQPAERGGQELRKEAEAFLEKGYLNEQVPAFTSTRS 60
DB      1 MWARVGLLRALQLLWGLHDAQPAERGGQELRKEAEAFLEKGYLNEQVPAFTSTRS 60
QY      61 DAIRAFQWQSQPVSGVLDRAFLRQWTRPRCCVDTJNSAAAEARISDLFAHRTWMRK 120
DB      61 DAIRAFQWQSQPVSGVLDRAFLRQWTRPRCCVDTJNSAAAEARISDLFAHRTWMRK 120
```

```
QY 121 KRFAKQGNKWKYKOHLSYRLVNMPEHLRSRQFAPCAPSSCGATSORMSGRPOAPLT 180
Db 121 KRFAKQGNKWKYKOHLSYRLVNMPEHLRSRQFAPCAPSSCGATSORMSGRPOAPLT 180
QY 181 SCSPSSKGTITGGMAMPMAWAQAPWRTPTFLPRRGEAHFDODERMSLSRRGRNLFVYLAH 240
Db 181 SCSPSSKGTITGGMAMPMAWAQAPWRTPTFLPRRGEAHFDODERMSLSRRGRNLFVYLAH 240
QY 241 EIGHTGLTHSPAPRALMAPYKRLGRDALLSMDVLAQVSLYKRPJGGSVAVOLPKLTF 300
Db 241 EIGHTGLTHSPAPRALMAPYKRLGRDALLSMDVLAQVSLYKRPJGGSVAVOLPKLTF 300
QY 301 TDFETWDSYSPQGRRPETQGPYCHSSFDATITVDROOQLYTFKSHFMEVAADGVSEPR 360
Db 301 TDFETWDSYSPQGRRPETQGPYCHSSFDATITVDROOQLYTFKSHFMEVAADGVSEPR 360
QY 361 PLOERWVGLPPIEAAVSLNDGDFYFFKGRGCRWFRGPKRWMLPOLCAAGI.PRHDA 420
Db 361 PLOERWVGLPPIEAAVSLNDGDFYFFKGRGCRWFRGPKRWMLPOLCAAGI.PRHDA 420
QY 421 ALFPPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGSIPEEVSALPRPDGSIIF 480
Db 421 ALFPPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGSIPEEVSALPRPDGSIIF 480
QY 481 RDDRWRMLDQAKLQATTSGRWATELPMWGMCHANSALP 520
Db 481 RDDRWRMLDQAKLQATTSGRWATELPMWGMCHANSALP 520
```

## RESULT 2

```
US-09-737-353-2
; Sequence 2, Application US/09737353
; Patent No. US20010036648A1
; GENERAL INFORMATION:
; APPLICANT: Christopher Donald Southan
; APPLICANT: Stephen Anthony Hughes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30222
; CURRENT APPLICATION NUMBER: US/09/737,353
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: 0001898.6
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FaSTSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-737-353-2
```

```
Query Match 87.2%; Score 2472; DB 9; Length 520;
Best Local Similarity 88.9%; Pred. No. 2,7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQILLMGHLDAQPAERGQELRKEAEAFLEKYGLNEQVPAAPTSTRFS 60
Db 1 MVARVGLLRALQILLMGHLDAQPAERGQELRKEAEAFLEKYGLNEQVPAAPTSTRFS 60
QY 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTPRCGVTDNTSYAAMERISDLFARHRTKMRK 120
Db 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTPRCGVTDNTSYAAMERISDLFARHRTKMRK 120
QY 121 KRFAKQGNKWKYKOHLSYRLVNMPEHLRSRQFAPCAPSSCGATSORMS--SGRQP 175
Db 121 KRFAKQGNKWKYKOHLSYRLVNMPEHLRSRQFAPCAPSSCGATSORMS--SGRQP 175
QY 176 QAPLTSGSPSS-----KGTITGGMAMPMAWAQAPWRTPTFLPRRGEAHFDODERMSLSRR 229
Db 176 QAPLTSGSPSS-----KGTITGGMAMPMAWAQAPWRTPTFLPRRGEAHFDODERMSLSRR 229
QY 229 EAPAT--GPADIRLTFQGDHNDLGNAPDGGALAAHFLPRRGEAHFDODERMSLSRR 229
Db 229 EAPAT--GPADIRLTFQGDHNDLGNAPDGGALAAHFLPRRGEAHFDODERMSLSRR 229
QY 230 RGRNLFVYLAHEIGHTGLTHSPAPRALMAPYKRLGRDALLSMDVLAQVSLYKRPJG 289
Db 230 RGRNLFVYLAHEIGHTGLTHSPAPRALMAPYKRLGRDALLSMDVLAQVSLYKRPJG 289
```

```
Db 230 RGRNLFVYLAHEIGHTGLTHSPAPRALMAPYKRLGRDALLSMDVLAQVSLYKRPJG 289
QY 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQGPYCHSSFDATITVDROOQLYTFKSHFWE 349
Db 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQGPYCHSSFDATITVDROOQLYTFKSHFWE 349
QY 350 VAADGVSEPRP.LQERWVGLPPIEAAVSLNDGDFYFFKGRGCRWFRGPKRWMLPOLC 409
Db 350 VAADGVSEPRP.LQERWVGLPPIEAAVSLNDGDFYFFKGRGCRWFRGPKRWMLPOLC 409
QY 410 RAGGLPRHPDALFFPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGSIPEEVSAGA 469
Db 410 RAGGLPRHPDALFFPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGSIPEEVSAGA 469
QY 470 LPRPDGSIIFPRDWRMLDQAKLQATTSGRWATELPMWGMCHANSALP 520
Db 470 LPRPDGSIIFPRDWRMLDQAKLQATTSGRWATELPMWGMCHANSALP 520
```

## RESULT 3

```
US-09-950-510-2
; Sequence 2, Application US/09950510
; Patent No. US20020150978A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, RORY
; TITLE OF INVENTION: 46798, A No. US20020150978A1 Human Matrix Metalloproteinase And
; FILE REFERENCE: 10147-4501
; CURRENT FILING DATE: US/09/950,510
; PRIOR APPLICATION NUMBER: 2001-09-10
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-510-2
```

```
Query Match 87.2%; Score 2472; DB 9; Length 520;
Best Local Similarity 88.9%; Pred. No. 2,7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQILLMGHLDAQPAERGQELRKEAEAFLEKYGLNEQVPAAPTSTRFS 60
Db 1 MVARVGLLRALQILLMGHLDAQPAERGQELRKEAEAFLEKYGLNEQVPAAPTSTRFS 60
QY 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTPRCGVTDNTSYAAMERISDLFARHRTKMRK 120
Db 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTPRCGVTDNTSYAAMERISDLFARHRTKMRK 120
QY 121 KRFAKQGNKWKYKOHLSYRLVNMPEHLRSRQFAPCAPSSCGATSORMS--SGRQP 175
Db 121 KRFAKQGNKWKYKOHLSYRLVNMPEHLRSRQFAPCAPSSCGATSORMS--SGRQP 175
QY 176 QAPLTSGSPSS-----KGTITGGMAMPMAWAQAPWRTPTFLPRRGEAHFDODERMSLSRR 229
Db 176 QAPLTSGSPSS-----KGTITGGMAMPMAWAQAPWRTPTFLPRRGEAHFDODERMSLSRR 229
QY 229 EAPAT--GPADIRLTFQGDHNDLGNAPDGGALAAHFLPRRGEAHFDODERMSLSRR 229
Db 229 EAPAT--GPADIRLTFQGDHNDLGNAPDGGALAAHFLPRRGEAHFDODERMSLSRR 229
QY 230 SVAVOLPGKLTDFETWDSYSPQGRRPETQGPYCHSSFDATITVDROOQLYTFKSHFWE 349
Db 230 SVAVOLPGKLTDFETWDSYSPQGRRPETQGPYCHSSFDATITVDROOQLYTFKSHFWE 349
QY 350 VAADGVSEPRP.LQERWVGLPPIEAAVSLNDGDFYFFKGRGCRWFRGPKRWMLPOLC 409
Db 350 VAADGVSEPRP.LQERWVGLPPIEAAVSLNDGDFYFFKGRGCRWFRGPKRWMLPOLC 409
QY 410 RAGGLPRHPDALFFPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGSIPEEVSAGA 469
Db 410 RAGGLPRHPDALFFPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGSIPEEVSAGA 469
```



```
Db      410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
Qy      470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
        |||
        470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520

RESULT 4
US-09-950-510-21
; Sequence 21, Application US/09950510
; Patent No. US20020150978A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory
; TITLE OF INVENTION: 46798, A No. US20020150978A1e1 Human Matrix Metalloproteinase And
; FILE REFERENCE: 10147-45U1
; CURRENT APPLICATION NUMBER: US/09/950,510
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/251,156
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 21
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-510-21

Query Match      87.2%; Score 2472; DB 9; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 4/2; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

Qy      1 MVARVGLLRALQLLLMGLDQPAERGGQELRKEAEAFLEKXGYLNEQVPKAPTSTRS 60
        |||
        1 MVARVGLLRALQLLLMGLDQPAERGGQELRKEAEAFLEKXGYLNEQVPKAPTSTRS 60
Db      61 DAIRAFQWVSQLPVSGVLDRAITLRQMTBRCGVTDTNSYAAMAEKISDLFAHRTKMRK 120
        |||
        61 DAIRAFQWVSQLPVSGVLDRAITLRQMTBRCGVTDTNSYAAMAEKISDLFAHRTKMRK 120
Db      121 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
        |||
        121 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
Db      122 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
        |||
        122 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
Qy      176 QAPLTSGSPSS-----KGTITMGWAMPPLMAOGAPWRTPELPRRGAHFDODERWSLSRR 229
        |||
        176 QAPLTSGSPSS-----KGTITMGWAMPPLMAOGAPWRTPELPRRGAHFDODERWSLSRR 229
Db      172 EAPAT--GPADIRLTFPGQDNDHDLGNAFDGCGALAHAFILPRRGAHFDODERWSLSRR 229
        |||
        172 EAPAT--GPADIRLTFPGQDNDHDLGNAFDGCGALAHAFILPRRGAHFDODERWSLSRR 229
Qy      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
        |||
        230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
Db      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
        |||
        230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
Qy      290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
        |||
        290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
Db      290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
        |||
        290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
Qy      350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
        |||
        350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
Db      350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
        |||
        350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
Qy      410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
        |||
        410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
Db      410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
        |||
        410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
Qy      470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
        |||
        470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
Db      470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
        |||
        470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520

RESULT 5
US-09-950-510-24
; Sequence 24, Application US/09950510
; Patent No. US20020150978A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory
; TITLE OF INVENTION: 46798, A No. US20020150978A1e1 Human Matrix Metalloproteinase And
; FILE REFERENCE: 10147-45U1
; CURRENT APPLICATION NUMBER: US/09/950,510
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/251,156
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 24
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-510-24

Query Match      87.2%; Score 2472; DB 9; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 4/2; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

Qy      1 MVARVGLLRALQLLLMGLDQPAERGGQELRKEAEAFLEKXGYLNEQVPKAPTSTRS 60
        |||
        1 MVARVGLLRALQLLLMGLDQPAERGGQELRKEAEAFLEKXGYLNEQVPKAPTSTRS 60
Db      61 DAIRAFQWVSQLPVSGVLDRAITLRQMTBRCGVTDTNSYAAMAEKISDLFAHRTKMRK 120
        |||
        61 DAIRAFQWVSQLPVSGVLDRAITLRQMTBRCGVTDTNSYAAMAEKISDLFAHRTKMRK 120
Db      121 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
        |||
        121 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
Db      122 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
        |||
        122 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
Qy      176 QAPLTSGSPSS-----KGTITMGWAMPPLMAOGAPWRTPELPRRGAHFDODERWSLSRR 229
        |||
        176 QAPLTSGSPSS-----KGTITMGWAMPPLMAOGAPWRTPELPRRGAHFDODERWSLSRR 229
Db      172 EAPAT--GPADIRLTFPGQDNDHDLGNAFDGCGALAHAFILPRRGAHFDODERWSLSRR 229
        |||
        172 EAPAT--GPADIRLTFPGQDNDHDLGNAFDGCGALAHAFILPRRGAHFDODERWSLSRR 229
Qy      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
        |||
        230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
Db      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
        |||
        230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
Qy      290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
        |||
        290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
Db      290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
        |||
        290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
Qy      350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
        |||
        350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
Db      350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
        |||
        350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
Qy      410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
        |||
        410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
Db      410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
        |||
        410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
Qy      470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
        |||
        470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
Db      470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
        |||
        470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520

RESULT 6
US-10-028-072-144
; Sequence 144, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
```

```
; APPLICANT: Curtis, Rory
; TITLE OF INVENTION: 46798, A No. US20020150978A1e1 Human Matrix Metalloproteinase And
; FILE REFERENCE: 10147-45U1
; CURRENT APPLICATION NUMBER: US/09/950,510
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/251,156
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 24
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-510-24

Query Match      87.2%; Score 2472; DB 9; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 4/2; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

Qy      1 MVARVGLLRALQLLLMGLDQPAERGGQELRKEAEAFLEKXGYLNEQVPKAPTSTRS 60
        |||
        1 MVARVGLLRALQLLLMGLDQPAERGGQELRKEAEAFLEKXGYLNEQVPKAPTSTRS 60
Db      61 DAIRAFQWVSQLPVSGVLDRAITLRQMTBRCGVTDTNSYAAMAEKISDLFAHRTKMRK 120
        |||
        61 DAIRAFQWVSQLPVSGVLDRAITLRQMTBRCGVTDTNSYAAMAEKISDLFAHRTKMRK 120
Db      121 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
        |||
        121 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
Db      122 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
        |||
        122 KRFPAQGNKMYKQHSYRLVNMPEHLRSRQFGAPCAPSPSCGATS---GRSPQ 175
Qy      176 QAPLTSGSPSS-----KGTITMGWAMPPLMAOGAPWRTPELPRRGAHFDODERWSLSRR 229
        |||
        176 QAPLTSGSPSS-----KGTITMGWAMPPLMAOGAPWRTPELPRRGAHFDODERWSLSRR 229
Db      172 EAPAT--GPADIRLTFPGQDNDHDLGNAFDGCGALAHAFILPRRGAHFDODERWSLSRR 229
        |||
        172 EAPAT--GPADIRLTFPGQDNDHDLGNAFDGCGALAHAFILPRRGAHFDODERWSLSRR 229
Qy      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
        |||
        230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
Db      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
        |||
        230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALISWDVLAVQSLYKGPLCG 289
Qy      290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
        |||
        290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
Db      290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
        |||
        290 SVAVOLPGKLFDPFETWDSYSPQGRPETQGRKCHSSFDATTVRQOOLYIFKSGHFE 349
Qy      350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
        |||
        350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
Db      350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
        |||
        350 VAADGNVSEPRPLQERWGLPNNIBAAVSLNDGDFEFYFKGRCRFRGPKVWGILPOLC 409
Qy      410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
        |||
        410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
Db      410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
        |||
        410 RAGGLPRHPDAAFPPEPLRRLILFKGARYVTLARGLQVEPYPRSLQDMGIPBEVSGA 469
Qy      470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
        |||
        470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
Db      470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
        |||
        470 LPRPGSIIFFRDDRYWRDLQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520

RESULT 6
US-10-028-072-144
; Sequence 144, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
```

APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang  
FILE OF INVENTION:  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/028,072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063227  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063229  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17

PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069212  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083222  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086414

```

/ PRIOR FILING DATE: 1998-05-22
/ PRIOR APPLICATION NUMBER: 60/086430
/ PRIOR FILING DATE: 1998-05-22
/ PRIOR APPLICATION NUMBER: 60/087106
/ PRIOR FILING DATE: 1998-05-28
/ PRIOR APPLICATION NUMBER: 60/088026
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088730
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088741
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088810
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088858
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/089532
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089599
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089907
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089947
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/090349
/ PRIOR FILING DATE: 1998-06-23
/ PRIOR APPLICATION NUMBER: 60/090429
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090445
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090538
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090663
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/091360
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091519
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091982
/ PRIOR FILING DATE: 1998-07-07

```

```

Query Match      87.2%; Score 2472; DB 14; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

```

```

QY 1 MVARVGLLRALQLLLMGLHDAQPARGGQELRKEAEFLERKGYLNEQVPAKPTSTRS 60
DB 1 MVARVGLLRALQLLLMGLHDAQPARGGQELRKEAEFLERKGYLNEQVPAKPTSTRS 60
QY 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDTNYSYAAMERISDLFAHRTQMRK 120
DB 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDTNYSYAAMERISDLFAHRTQMRK 120
QY 121 KRPFAQGNKMYKQHLSTYRLVNMPEHLRSGQFAPCAPSPSSCGATS--GRMS--SGRPOP 175
DB 121 KRPFAQGNKMYKQHLSTYRLVNMPEHLRSGQFAPCAPSPSSCGATS--GRMS--SGRPOP 175
QY 176 QAPLTSGSPSS-----KGTMMGAMPLMAQGAAPRTPLPRRGAHFDDERMSLSR 229
DB 176 QAPLTSGSPSS-----KGTMMGAMPLMAQGAAPRTPLPRRGAHFDDERMSLSR 229
QY 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYKGPLG 289
DB 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYKGPLG 289
QY 290 SVAVOLPGKLFDTFETWDSYSPQGRRPETQGRPKYCHSSFDALITVDROOOLYIFKSGHFW 349
DB 290 SVAVOLPGKLFDTFETWDSYSPQGRRPETQGRPKYCHSSFDALITVDROOOLYIFKSGHFW 349
QY 350 VAADGNVSEBRPLQERWVGLPNIIEAAVSLNDGDYFFPKGRCWCFRGPCKVWGLPOLC 409
DB 350 VAADGNVSEBRPLQERWVGLPNIIEAAVSLNDGDYFFPKGRCWCFRGPCKVWGLPOLC 409
QY 410 RAGGLPRHDDALFPPLRLRLIFKGARVYVLAAGGLQVEPYPRSLQDWGIPREVSQA 469

```

```

DB 410 RAGGLPRHDDALFPPLRLRLIFKGARVYVLAAGGLQVEPYPRSLQDWGIPREVSQA 469
QY 470 LRPDGSITFFRDDRWRMDQAKLQATTSGRVATELPMWGCYHANSQSALF 520
DB 470 LRPDGSITFFRDDRWRMDQAKLQATTSGRVATELPMWGCYHANSQSALF 520

```

# RESULT 7

```

US-10-140-808-144
/ Sequence 144, Application US/10140808
/ Publication No. US20030017563A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: Deforge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330R1C182
/ CURRENT APPLICATION NUMBER: US/10/140,808
/ CURRENT FILING DATE: 2002-05-07
/ Prior Application removed - See file Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 144
/ LENGTH: 520
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-140-808-144

```

```

Query Match      87.2%; Score 2472; DB 14; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

```

```

QY 1 MVARVGLLRALQLLLMGLHDAQPARGGQELRKEAEFLERKGYLNEQVPAKPTSTRS 60
DB 1 MVARVGLLRALQLLLMGLHDAQPARGGQELRKEAEFLERKGYLNEQVPAKPTSTRS 60
QY 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDTNYSYAAMERISDLFAHRTQMRK 120
DB 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDTNYSYAAMERISDLFAHRTQMRK 120
QY 121 KRPFAQGNKMYKQHLSTYRLVNMPEHLRSGQFAPCAPSPSSCGATS--GRMS--SGRPOP 175
DB 121 KRPFAQGNKMYKQHLSTYRLVNMPEHLRSGQFAPCAPSPSSCGATS--GRMS--SGRPOP 175
QY 176 QAPLTSGSPSS-----KGTMMGAMPLMAQGAAPRTPLPRRGAHFDDERMSLSR 229
DB 176 QAPLTSGSPSS-----KGTMMGAMPLMAQGAAPRTPLPRRGAHFDDERMSLSR 229
QY 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYKGPLG 289
DB 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYKGPLG 289
QY 290 SVAVOLPGKLFDTFETWDSYSPQGRRPETQGRPKYCHSSFDALITVDROOOLYIFKSGHFW 349
DB 290 SVAVOLPGKLFDTFETWDSYSPQGRRPETQGRPKYCHSSFDALITVDROOOLYIFKSGHFW 349
QY 350 VAADGNVSEBRPLQERWVGLPNIIEAAVSLNDGDYFFPKGRCWCFRGPCKVWGLPOLC 409

```

Db 350 VAADNVSEBRPLOERWGLPNIIEAAVSLNDGFFYFKGRCWRFRGPKVWGLPOLC 409  
Qy 410 RAGGLPRHPDALFPPPLRLILFKGARYYVLARGLQVEBPYPSLQDWGSIPEEVSQA 469  
Db 410 RAGGLPRHPDALFPPPLRLILFKGARYYVLARGLQVEBPYPSLQDWGSIPEEVSQA 469  
Qy 470 LPRPDGSIIFPRDRYWRDLQAKLQATTSGRWATELPMWGCWANHSGSALF 520  
Db 470 LPRPDGSIIFPRDRYWRDLQAKLQATTSGRWATELPMWGCWANHSGSALF 520

## RESULT 8

US-10-121-049-144  
Sequence 144, Application US/10121049  
Publication No. US2003002239A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Geo, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C17  
CURRENT FILING DATE: 2002-04-12  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 144  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-121-049-144

Query Match 87.2%; Score 2472; DB 14; Length 520;  
Best Local Similarity 88.9%; Pred. No. 2,7e-203;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

Qy 1 MVARVGLLLRALQLLMGLHDAQPAERGGQELRKEAEFLKRYGYLNEQVPAPTSTRFS 60  
Db 1 MVARVGLLLRALQLLMGLHDAQPAERGGQELRKEAEFLKRYGYLNEQVPAPTSTRFS 60  
Qy 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTFRPGCVTDJNSYAAMAEISDLFAHRTTKMRK 120  
Db 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTFRPGCVTDJNSYAAMAEISDLFAHRTTKMRK 120  
Qy 121 KRPAQGNKWKYKQHLSTYRLVWMPBEHLRSRQFCAPCAPSSCGATS---QRMS--SGRPP 175  
Db 121 KRPAQGNKWKYKQHLSTYRLVWMPBEHLRSRQFCAPCAPSSCGATS---QRMS--SGRPP 175  
Qy 176 QAPLTSGSPSS-----KGTITMGAMPMLMAQGAPEWPTPLPRRGEAHFDODERWSLSRR 229  
Db 172 EAPAT--GPADIRLTFQGDHNDGLGNAFDGPGALAHAFPRRGEAHFDODERWSLSRR 229  
Qy 230 RGRNLFVVLAAHEIGTTLGLTHSPAPRALMAPIYKRLGRDALLSMDVLAQSLYKPKLGG 289  
Db 230 RGRNLFVVLAAHEIGTTLGLTHSPAPRALMAPIYKRLGRDALLSMDVLAQSLYKPKLGG 289  
Qy 290 SVAVQLPGLKFTDEPTWDSYSPQGRPEQGPYKCHSSPDAITVDROQOLYIFKQSHFWE 349  
Db 290 SVAVQLPGLKFTDEPTWDSYSPQGRPEQGPYKCHSSPDAITVDROQOLYIFKQSHFWE 349

Qy 350 VAADNVSEBRPLOERWGLPNIIEAAVSLNDGFFYFKGRCWRFRGPKVWGLPOLC 409  
Db 350 VAADNVSEBRPLOERWGLPNIIEAAVSLNDGFFYFKGRCWRFRGPKVWGLPOLC 409  
Qy 410 RAGGLPRHPDALFPPPLRLILFKGARYYVLARGLQVEBPYPSLQDWGSIPEEVSQA 469  
Db 410 RAGGLPRHPDALFPPPLRLILFKGARYYVLARGLQVEBPYPSLQDWGSIPEEVSQA 469  
Qy 470 LPRPDGSIIFPRDRYWRDLQAKLQATTSGRWATELPMWGCWANHSGSALF 520  
Db 470 LPRPDGSIIFPRDRYWRDLQAKLQATTSGRWATELPMWGCWANHSGSALF 520

## RESULT 9

US-10-123-904-144  
Sequence 144, Application US/10123904  
Publication No. US2003002238A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Geo, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C54  
CURRENT FILING DATE: 2002-04-16  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 144  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-123-904-144

Query Match 87.2%; Score 2472; DB 14; Length 520;  
Best Local Similarity 88.9%; Pred. No. 2,7e-203;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

Qy 1 MVARVGLLLRALQLLMGLHDAQPAERGGQELRKEAEFLKRYGYLNEQVPAPTSTRFS 60  
Db 1 MVARVGLLLRALQLLMGLHDAQPAERGGQELRKEAEFLKRYGYLNEQVPAPTSTRFS 60  
Qy 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTFRPGCVTDJNSYAAMAEISDLFAHRTTKMRK 120  
Db 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTFRPGCVTDJNSYAAMAEISDLFAHRTTKMRK 120  
Qy 121 KRPAQGNKWKYKQHLSTYRLVWMPBEHLRSRQFCAPCAPSSCGATS---QRMS--SGRPP 175  
Db 121 KRPAQGNKWKYKQHLSTYRLVWMPBEHLRSRQFCAPCAPSSCGATS---QRMS--SGRPP 175  
Qy 176 QAPLTSGSPSS-----KGTITMGAMPMLMAQGAPEWPTPLPRRGEAHFDODERWSLSRR 229  
Db 172 EAPAT--GPADIRLTFQGDHNDGLGNAFDGPGALAHAFPRRGEAHFDODERWSLSRR 229  
Qy 230 RGRNLFVVLAAHEIGTTLGLTHSPAPRALMAPIYKRLGRDALLSMDVLAQSLYKPKLGG 289  
Db 230 RGRNLFVVLAAHEIGTTLGLTHSPAPRALMAPIYKRLGRDALLSMDVLAQSLYKPKLGG 289

```

QY 290 SVAVOLPGKLTDFDFTWDSYSPQGRRPETQGRKYCHSSFDALTVDROOQLYIFKSGHFW 349
DB 290 SVAVOLPGKLTDFDFTWDSYSPQGRRPETQGRKYCHSSFDALTVDROOQLYIFKSGHFW 349
QY 350 VAADGVNSRPRLQERWVGLPPIIEAAVSLNDGDFYFFKGRGRCWRFRGPKVWGLPQLC 409
DB 350 VAADGVNSRPRLQERWVGLPPIIEAAVSLNDGDFYFFKGRGRCWRFRGPKVWGLPQLC 409
QY 410 RAGGLPRHDDALFFPPLRLILFKGARYYVLARGLQVEPYPRSLQDWGIPREVSGA 469
DB 410 RAGGLPRHDDALFFPPLRLILFKGARYYVLARGLQVEPYPRSLQDWGIPREVSGA 469
QY 470 LRPDGSIIFFRDDRYRLDQAKLQATTSGRWATELPMWGMCHANSGLALF 520
DB 470 LRPDGSIIFFRDDRYRLDQAKLQATTSGRWATELPMWGMCHANSGLALF 520

RESULT 10
US-10-140-470-144
; Sequence 144, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 144
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-144

Query Match 87.2%; Score 2472; DB 14; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

```

```

DB 230 RGRNLFVVLAAHIGTLGLTHSPAPRALMAPYKRLGRDALLSWDVLAVQSLYXKPLGG 289
QY 290 SVAVOLPGKLTDFDFTWDSYSPQGRRPETQGRKYCHSSFDALTVDROOQLYIFKSGHFW 349
DB 290 SVAVOLPGKLTDFDFTWDSYSPQGRRPETQGRKYCHSSFDALTVDROOQLYIFKSGHFW 349
QY 350 VAADGVNSRPRLQERWVGLPPIIEAAVSLNDGDFYFFKGRGRCWRFRGPKVWGLPQLC 409
DB 350 VAADGVNSRPRLQERWVGLPPIIEAAVSLNDGDFYFFKGRGRCWRFRGPKVWGLPQLC 409
QY 410 RAGGLPRHDDALFFPPLRLILFKGARYYVLARGLQVEPYPRSLQDWGIPREVSGA 469
DB 410 RAGGLPRHDDALFFPPLRLILFKGARYYVLARGLQVEPYPRSLQDWGIPREVSGA 469
QY 470 LRPDGSIIFFRDDRYRLDQAKLQATTSGRWATELPMWGMCHANSGLALF 520
DB 470 LRPDGSIIFFRDDRYRLDQAKLQATTSGRWATELPMWGMCHANSGLALF 520

RESULT 11
US-10-175-746-144
; Sequence 144, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 144
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-144

Query Match 87.2%; Score 2472; DB 14; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

```

```
Db 172 EAPAT--GPADIRLTFPGDNDGLGNAFDGPGALAHAFPRRGEAHFPDDEKMSLSRR 229
Qy 230 RGRNLFVYLAHEIGTTLGLTSPAPRALMAPYKRLGSDALLSMDVLAVOGLYKPLGG 289
Db 230 RGRNLFVYLAHEIGTTLGLTSPAPRALMAPYKRLGSDALLSMDVLAVOGLYKPLGG 289
Qy 290 SVAVOLPGKLTFTDEFTWDSYSPQGRRPETQPKYCHSFPDAITYDROOQLYIFKGSHPWE 349
Db 290 SVAVOLPGKLTFTDEFTWDSYSPQGRRPETQPKYCHSFPDAITYDROOQLYIFKGSHPWE 349
Qy 350 VAAOENVSEPRPLDERWVGLPNNIEAAVSLNDGDFYFFKGCRCMRFRGPKPVWGLPOLC 409
Db 350 VAAOENVSEPRPLDERWVGLPNNIEAAVSLNDGDFYFFKGCRCMRFRGPKPVWGLPOLC 409
Qy 410 RAGGLPRHPDAALEFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGCIPEEVSQA 469
Db 410 RAGGLPRHPDAALEFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGCIPEEVSQA 469
Qy 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGCWANSALSALF 520
Db 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGCWANSALSALF 520
```

```
RESULT 12
US-10-176-918-144
; Sequence 144, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; Prior Filing Date: 2002-06-20
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 144
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-918-144
```

```
Query Match 87.2%; Score 2472; DB 14; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;
Qy 1 MVARVGLLRALQLLWGHLDLDAQPAERGGQELRKEAEAFLEKYGLNEQVPAKPTSTRFS 60
Db 1 MVARVGLLRALQLLWGHLDLDAQPAERGGQELRKEAEAFLEKYGLNEQVPAKPTSTRFS 60
Qy 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDITNSYAAMABRISDLFARHRTKMRK 120
Db 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDITNSYAAMABRISDLFARHRTKMRK 120
Qy 121 KRFAQGNKMYKQHLSTYRLVNWPEHLRSQFGAPCAPSSCGATS---QRMS--SGRPOP 175
Db 121 KRFAQGNKMYKQHLSTYRLVNWPEHLRSQFGAPCAPSSCGATS---QRMS--SGRPOP 175
```

```
Qy 176 QAPLTSGSFS-----KQTTMGWAMPLMQAGAWRTPTFLPRRGEAHFPDDEKMSLSRR 229
Db 172 EAPAT--GPADIRLTFPGDNDGLGNAFDGPGALAHAFPRRGEAHFPDDEKMSLSRR 229
Qy 230 RGRNLFVYLAHEIGTTLGLTSPAPRALMAPYKRLGSDALLSMDVLAVOGLYKPLGG 289
Db 230 RGRNLFVYLAHEIGTTLGLTSPAPRALMAPYKRLGSDALLSMDVLAVOGLYKPLGG 289
Qy 290 SVAVOLPGKLTFTDEFTWDSYSPQGRRPETQPKYCHSFPDAITYDROOQLYIFKGSHPWE 349
Db 290 SVAVOLPGKLTFTDEFTWDSYSPQGRRPETQPKYCHSFPDAITYDROOQLYIFKGSHPWE 349
Qy 350 VAAOENVSEPRPLDERWVGLPNNIEAAVSLNDGDFYFFKGCRCMRFRGPKPVWGLPOLC 409
Db 350 VAAOENVSEPRPLDERWVGLPNNIEAAVSLNDGDFYFFKGCRCMRFRGPKPVWGLPOLC 409
Qy 410 RAGGLPRHPDAALEFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGCIPEEVSQA 469
Db 410 RAGGLPRHPDAALEFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGCIPEEVSQA 469
Qy 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGCWANSALSALF 520
Db 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGCWANSALSALF 520
```

```
RESULT 13
US-10-176-921-144
; Sequence 144, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C286
; CURRENT APPLICATION NUMBER: US/10/176,921
; Prior Filing Date: 2002-06-20
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 144
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-921-144
```

```
Query Match 87.2%; Score 2472; DB 14; Length 520;
Best Local Similarity 88.9%; Pred. No. 2.7e-203;
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;
Qy 1 MVARVGLLRALQLLWGHLDLDAQPAERGGQELRKEAEAFLEKYGLNEQVPAKPTSTRFS 60
Db 1 MVARVGLLRALQLLWGHLDLDAQPAERGGQELRKEAEAFLEKYGLNEQVPAKPTSTRFS 60
Qy 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDITNSYAAMABRISDLFARHRTKMRK 120
Db 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDITNSYAAMABRISDLFARHRTKMRK 120
```

|    |     |  |      |    |        |     |
|----|-----|--|------|----|--------|-----|
| Qy | 121 | KRFAQGNKMYQXOHLSTRVWNPENHRSQFBCAPSPSCGTS---                  | QWMS | -- | SGRPOP | 175 |
| Db | 121 | KRFAGQGNKMYQXOHLSTRVWNPENH---                                |      |    |        |     |
| Qy | 121 | KRFAGQGNKMYQXOHLSTRVWNPENH---                                |      |    |        |     |
| Db | 121 | KRFAGQGNKMYQXOHLSTRVWNPENH---                                |      |    |        |     |
| Qy | 176 | QAPLTSGPSS-----KGTTWGMAPMLAQAPWRTPFLPRRGEAHFQDDBRWSLSRR      |      |    |        | 229 |
| Db | 172 | EAPAT--GPADIRLTFQGDHNDGIDGNAFDGPGCALAHAVLPRRGEAHFQDDBRWSLSRR |      |    |        | 229 |
| Qy | 230 | RGRLFYVTLAHEIGTTLGLTTHSPAPALMAFYKRLGRDALLSWDVLAVQSLYGRPLGG   |      |    |        | 289 |
| Db | 230 | RGRLFYVTLAHEIGTTLGLTTHSPAPALMAFYKRLGRDALLSWDVLAVQSLYGRPLGG   |      |    |        | 289 |
| Qy | 290 | SVAVOLPGKLFNDFETWDSYSPQGRREPMQGPXYCHSSPDAIVVDBOQOLYIFKSGHFW  |      |    |        | 349 |
| Db | 290 | SVAVOLPGKLFNDFETWDSYSPQGRREPMQGPXYCHSSPDAIVVDBOQOLYIFKSGHFW  |      |    |        | 349 |
| Qy | 350 | VAADNVSEPRLOERWGLPNIIEAAVSLNDGDYFFFGGRGCMRFRGPKPMGLPOLC      |      |    |        | 409 |
| Db | 350 | VAADNVSEPRLOERWGLPNIIEAAVSLNDGDYFFFGGRGCMRFRGPKPMGLPOLC      |      |    |        | 409 |
| Qy | 410 | RAGGLPRHPDALFEPPLRRLILFKGARYYVLARGLOVEPYPRLQDWGGLPEEVSQA     |      |    |        | 469 |
| Db | 410 | RAGGLPRHPDALFEPPLRRLILFKGARYYVLARGLOVEPYPRLQDWGGLPEEVSQA     |      |    |        | 469 |
| Qy | 470 | LPRPGSIIFFRDDRYWRLDQAKLQATTSGRNATELPMWGCMHANSGLAF            |      |    |        | 520 |
| Db | 470 | LPRPGSIIFFRDDRYWRLDQAKLQATTSGRNATELPMWGCMHANSGLAF            |      |    |        | 520 |

RESULT 14  
US-10-137-865-144  
; Sequence 144, Application US/10137865  
; Publication No. US20030032155A1

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C154
CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 144
LENGTH: 520
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-865-144

```

|                           |       |                    |           |            |
|---------------------------|-------|--------------------|-----------|------------|
| Query Match               | 87.2% | Score 2472         | DB 14     | Length 520 |
| Best Local Similarity     | 88.9% | Pred. No. 2.7e-203 |           |            |
| Matches 472; Conservative | 4     | Mismatches 33      | Indels 22 | Gaps 5     |

Qy 1 MVARVGLILRLALQILLMGHLDPQAPBRGGQELRKEAEALBKGYGYNEQVPKAP15TRRS 60

Dy 1 MVARVGLILRLALQILLMGHLDPQAPBRGGQELRKEAEALBKGYGYNEQVPKAP15TRRS 60

Qy 61 DAIRAFQWVSQIPVSGVLDRAFLRLQMTTRRCGVTDPTNLSAANAERISDLPARRRTCKRRK 120

```

Db      61 DAIRPQWVSQLPVSGVLDRALRLQMTPRCCVTTNSTAAAEKRLSDLPFARHTQMRBK 120
Qy      121 KRFAKQGNKWKYKQHLSTYRLVNMPEHLRSRQFGCAPCPASSCGATS--GRWS--SGRPOP 175
Db      121 KRFAKQGNKWKYKQHLSTYRLVNMPEHL-----PEPAVGAQVRAAQLMSVSLBFW 171
Qy      176 QAPLTSSPSS-----KGTTMGMAFPLMAGAQWRTPFLPRRGAAHADOBERMLSRR 229
Db      172 EAPAT--GRADIRLTFPQGDHNDGIGNAFDGPQALAHAFLLRRGAAHPDOBERMLSRR 229
Qy      230 RGRNLFVYLAHEIGTTLG/THSPAPRALMAPYKRLGRDALLSMDVLAVOSLYGKPLG 289
Db      230 RGRNLFVYLAHEIGTTLG/THSPAPRALMAPYKRLGRDALLSMDVLAVOSLYGKPLG 289
Qy      290 SVAVOLPEKLTFTDFETWDSYSPQGRRPETQGRKYCHSFDATITVDROOQLYFKSGHFW 349
Db      290 SVAVOLPEKLTFTDFETWDSYSPQGRRPETQGRKYCHSFDATITVDROOQLYFKSGHFW 349
Qy      350 VAADGNVSEPRPLOERWGLPNNIEAALVSLNDGJFFPKGGRCMRFGRPKVWGLPOLC 409
Db      350 VAADGNVSEPRPLOERWGLPNNIEAALVSLNDGJFFPKGGRCMRFGRPKVWGLPOLC 409
Qy      410 RAGGIRPRPDALFPPRLRLLIFGARYYTLARGLQVPRPYPPSLQDWGSIPEVSGA 469
Db      410 RAGGIRPRPDALFPPRLRLLIFGARYYTLARGLQVPRPYPPSLQDWGSIPEVSGA 469
Qy      470 LPRPDGSIIFPRDDRYWRLODAKLOATTSGRVATELPMWGCNHAUSGSLAF 520
Db      470 LPRPDGSIIFPRDDRYWRLODAKLOATTSGRVATELPMWGCNHAUSGSLAF 520

```

RESULT 15  
US-10-140-474-144  
; Sequence 144, Application US/10140474  
; Publication No. US20030032156A1

```

1  APPLICANT: Beresini, Maureen
2  APPLICANT: Defoige, Laura
3  APPLICANT: Desnoyers, Luc
4  APPLICANT: Filvaroff, Ellen
5  APPLICANT: Gao, Wei-Qiang
6  APPLICANT: Gerritsen, Mary E.
7  APPLICANT: Goddard, Audrey
8  APPLICANT: Godowski, Paul J.
9  APPLICANT: Gurney, Austin L.
10 APPLICANT: Sherwood, Steven
11 APPLICANT: Smith, Victoria
12 APPLICANT: Stewart, Timothy A.
13 APPLICANT: Tumas, Daniel
14 APPLICANT: Watanabe, Colin K
15 APPLICANT: Wood, William
16 APPLICANT: Zhang, Zemin
17 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
18 FILE REFERENCE: P330R1C162
19 CURRENT APPLICATION NUMBER: US/10/140,474
20 CURRENT FILING DATE: 2002-05-06
21 Prior Application removed - See Palm or File Wrapper
22 NUMBER OF SEQ ID NOS: 550
23 SEQ ID NO 144
24 LENGTH: 520
25 TYPE: PRT
26 ORGANISM: Homo Sapien
27 US-10-140-474-144

```

|                           |        |                    |            |             |
|---------------------------|--------|--------------------|------------|-------------|
| Query Match               | 87.2%; | Score 2472;        | DB 14;     | Length 520; |
| Best Local Similarly      | 88.9%; | Pred. No. 2.7e203; |            |             |
| Matches 472; Conservative | 4;     | Mismatches 33;     | Indels 22; | Gaps 5      |

QY 1 MVARGLLRRLQILLMGLHDQPAERGGQELRKEAEAFLEKYGYLNEQVPKAPSTRFS 600

```
Db      1 MVAEVLRLALQILLWGHLDAPAEERGGEELRKEAEFLKYGYLEQVYKAPJSTRPS 60
Qy      61 DAIRAFQWVSQLPVSGVLDRAJLROMTRPRCGVTDJNSYAAMAEISDLFARHRTKMRK 120
Db      61 DAIRAFQWVSQLPVSGVLDRAJLROMTRPRCGVTDJNSYAAMAEISDLFARHRTKMRK 120
Qy      121 KRFAKQGNKMYKQHLSTRLVNWPBHLRSROFGAPCAPSSCGATS--QRWS--SGRPOP 175
Db      121 KRFAKQGNKMYKQHLSTRLVNWPBHLRSROFGAPCAPSSCGATS--QRWS--SGRPOP 171
Qy      176 QABLTSGSPSS-----KGTITGMWMPFLMAOGAPWRTPELPRRGEAHFDQDERWLSRR 229
Db      172 EADAT--GPADILTLTFPGDHNDGLGNAFPGGALAHAFLEPRRGEAHFDQDERWLSRR 229
Qy      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALSWDDVLAQSLYGKPLGG 289
Db      230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALSWDDVLAQSLYGKPLGG 289
Qy      290 SVAVOLPGKLFETDFTWDSYSPQGRRPETQGPYCHSSFDATTVDRQOOLYTFKSHFWE 349
Db      290 SVAVOLPGKLFETDFTWDSYSPQGRRPETQGPYCHSSFDATTVDRQOOLYTFKSHFWE 349
Qy      350 VAADGNVSEBRPLQERWVGLPVIIEAAYSLNDGDFYFFKGRGCRWFRGPKFVWGLPOLC 409
Db      350 VAADGNVSEBRPLQERWVGLPVIIEAAYSLNDGDFYFFKGRGCRWFRGPKFVWGLPOLC 409
Qy      410 RAGGLPRHPDALFPPLRLILFKGARYYVLARGGLOVEPYPRSLQDWGIPREVSQA 469
Db      410 RAGGLPRHPDALFPPLRLILFKGARYYVLARGGLOVEPYPRSLQDWGIPREVSQA 469
Qy      470 LPRPDGSIIFRRDDRYWRLDQAKLOATTSGRWATELPMWGCWHANSGSALP 520
Db      470 LPRPDGSIIFRRDDRYWRLDQAKLOATTSGRWATELPMWGCWHANSGSALP 520
```

Search completed: June 13, 2005, 18:15:10  
Job time : 159 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using bw model

Run on: June 13, 2005, 17:24:41 ; Search time 42 Seconds  
(without alignments)  
1191.255 Million cell updates/sec

Title: US-10-791-980-6  
Perfect score: 2834  
Sequence: 1 MVARVGLLRALLQLLWGH.....WATELPMWGMWANSALP 520

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 546   | 19.3        | 508    | 2 JCS082 | matrix metalloprot  |
| 2          | 534   | 18.8        | 669    | 2 I38029 | matrix metalloprot  |
| 3          | 517.5 | 18.3        | 582    | 2 I64471 | matrix metalloprot  |
| 4          | 514.5 | 18.2        | 582    | 2 I38028 | matrix metalloprot  |
| 5          | 466.5 | 16.5        | 492    | 2 A44399 | stromelysin 3 (EC   |
| 6          | 463.5 | 16.4        | 582    | 2 I48673 | matrix metalloprot  |
| 7          | 463   | 16.3        | 491    | 2 JCG197 | stromelysin 3 (EC   |
| 8          | 455.5 | 16.1        | 480    | 2 S13423 | stromelysin 3 (EC   |
| 9          | 448.5 | 15.8        | 477    | 1 KCMSL  | stromelysin 1 (EC   |
| 10         | 427   | 15.1        | 475    | 1 KCRTH  | stromelysin 1 (EC   |
| 11         | 426.5 | 15.0        | 471    | 2 A53711 | collagenase 3 (EC   |
| 12         | 425.5 | 15.0        | 476    | 1 KCRYS2 | stromelysin 2 (EC   |
| 13         | 410   | 14.5        | 483    | 2 JCS743 | matrix metalloprot  |
| 14         | 407.5 | 14.4        | 466    | 2 A23685 | interstitial colla  |
| 15         | 406   | 14.3        | 476    | 1 UC6505 | stromelysin 2 (EC   |
| 16         | 401   | 14.1        | 469    | 1 KCHU1  | interstitial colla  |
| 17         | 398   | 14.0        | 477    | 1 KCHUS1 | stromelysin 1 (EC   |
| 18         | 392   | 13.8        | 476    | 1 KCHUS2 | stromelysin 2 (EC   |
| 19         | 391.5 | 13.8        | 470    | 2 A49499 | metalloelastase HM  |
| 20         | 391   | 13.8        | 478    | 1 KCRBS1 | stromelysin 1 (EC   |
| 21         | 388.5 | 13.7        | 472    | 2 S29243 | interstitial colla  |
| 22         | 386   | 13.6        | 468    | 1 KCRBI  | interstitial colla  |
| 23         | 378   | 13.3        | 477    | 1 I51645 | stromelysin 3 (EC   |
| 24         | 368.5 | 13.0        | 469    | 1 KCPGI  | interstitial colla  |
| 25         | 366   | 12.9        | 462    | 2 A42401 | macrophage elastas  |
| 26         | 355.5 | 12.5        | 587    | 2 S41409 | enkephalin (EC 3.4. |
| 27         | 353.5 | 12.5        | 587    | 2 S12805 | enkephalin (EC 3.4. |
| 28         | 345.5 | 12.2        | 663    | 1 S46492 | gelatinase A (EC 3  |
| 29         | 344.5 | 12.2        | 467    | 1 KCHUN  | neutrophil collase  |

|    |       |      |     |           |                    |
|----|-------|------|-----|-----------|--------------------|
| 30 | 340.5 | 12.0 | 469 | 1 KCB01   | interstitial colla |
| 31 | 336.5 | 11.9 | 662 | 2 S70365  | gelatinase A (EC 3 |
| 32 | 322.5 | 11.4 | 660 | 3 A28153  | gelatinase A (EC 3 |
| 33 | 322   | 11.4 | 662 | 2 S34780  | gelatinase A (EC 3 |
| 34 | 318.5 | 11.2 | 712 | 1 I46031  | gelatinase B (EC 3 |
| 35 | 316.5 | 11.2 | 707 | 1 A53796  | gelatinase B (EC 3 |
| 36 | 313   | 11.0 | 662 | 2 A42496  | gelatinase A (EC 3 |
| 37 | 307   | 10.8 | 730 | 1 I52580  | gelatinase B (EC 3 |
| 38 | 307   | 10.8 | 730 | 2 JCG1456 | gelatinase B (EC 3 |
| 39 | 303.5 | 10.7 | 616 | 2 JCG7776 | matrix metalloprot |
| 40 | 301.5 | 10.6 | 707 | 1 A34458  | gelatinase B (EC 3 |
| 41 | 293   | 10.3 | 708 | 2 UC4364  | gelatinase B (EC 3 |
| 42 | 291.5 | 10.3 | 579 | 2 T37248  | probable matrix me |
| 43 | 291.5 | 10.3 | 598 | 2 T32166  | hypothetical prote |
| 44 | 278   | 9.8  | 521 | 2 T37252  | probable matrix me |
| 45 | 278   | 9.8  | 708 | 2 S62907  | gelatinase B (EC 3 |

ALIGNMENTS

RESULT 1

JCS082  
matrix metalloproteinase 18 (EC 3.4.24.-) precursor - human  
C:Species: Homo sapiens (man)  
C>Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 09-Jul-2004  
C:Accession: JCS082  
R:Crossin, J., Dudgeon, T.J., Catlin, G., Gearing, A.J.H., Clements, J.M.  
Biochem. Biophys. Res. Commun. 228, 494-498, 1996  
A>Title: Identification of MMP-18, a putative novel human matrix metalloproteinase.  
A:Reference number: JCS082; MUID:97079209; PMID:8920941  
A:Accession: JCS082  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-508 <COS>  
A:Cross-references: UNIPROT:Q99542; GB:Y08622  
A:Comment: This enzyme is involved in organ morphogenesis, embryonic development and pat  
sion and metastasis.  
C:Genetic8:  
A:Gene: GDB:IMPI8  
A:Cross-references: GDB:5584569; OMIM:601807  
C:Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloprotei  
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:53-256/Domain: matrix metalloproteinase homology <MMP>  
F:283-472/Domain: hemopexin repeat homology <PHX>  
F:185,212,216,222/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
F:212,216,222/Binding site: zinc, catalytic (His) (active) #status predicted  
F:213/Active site: Glu #status predicted

|                       |       |  |   |                                    |
|-----------------------|-------|--|---|------------------------------------|
| Query Match           | 19.3% | Score 546  | DB 2  | Length 508                         |
| Best Local Similarity | 31.4% | Pred. No. 1.5e-33  |   |                                    |
| Matches               | 158   | Conservative   | 75  | Mismatches 178; Indels 92; Gaps 19 |
| QY                    | 39    | FLKRYGLNEQVPAAPSTRF-----                                   | SDAIRAFQWVSQLPVSGVLRATLTROMTRRCGV           | 93                                 |
| DB                    | 31    | YLSGVYQKPL-----  | BGSNNFKPEDITFALRAFQEASLPSVSGQDDATRAMRQPRCGI | 87                                 |
| QY                    | 94    | TDVTSYAAMERISDLFPAHRTKRRKRRKPAKQGNKTKQHLSTRLVNMPHLASRQGA   | 153   |                                    |
| DB                    | 88    | EDPFR-----   | QKTLKYLILG-RMRKKHLTRILNLPSTL-----           | 120                                |
| QY                    | 154   | PCAPSSSCGA-----  | TSQKSSGRPOQAPLT-----                        | 194                                |
| DB                    | 121   | ---PPTRRALRLRAFQDMSN-----                                  | VAPLTFQEVQAGADIRLSFRGRSSYCSNTDFG            | 172                                |
| QY                    | 195   | AMPLMAQAPWRTPLPRGSAHPDODERWLSRRGRNLFVLAHEIGTLGLTHSPAP      | 254   |                                    |
| DB                    | 173   | PGRVLAHND-----   | IPELGSHVHDEDEFTGTYGVNLRIIAAHEVGHALGLGHSRYS  | 226                                |
| QY                    | 255   | RALMAFYTKRGKDALMSMDVLAQSLYGRPLGGSVAVALPGKLTDFDETWDSDYSPQGR | 314   |                                    |
| DB                    | 227   | QALMAFVEGYRPHFKLHPDDVAGIQALYGKK-----                       | SPVIRDEBEETELPT---VPPVPT                    | 280                                |

QY 315 RPEFGGRKCHSSPDATTVDRQOOLYIFKGSHEWFEYADGVNSEPRFL---QERWVGLPP 3711

Db 281 EPSMW-EDPSSSELDAWMLGPRKGTAFKGDYVWTVYSDSG---FGPLFRVYSALMEGLPG 3355

QY 372 NIEAAVSLNDGDFFFEFKGRCWRFGRGPRVWGLPOLCRAGGLPRHPDALFPPDLRLI 4311

Db 336 NLDAAYVSPRTQWIIHFPKDKVWRVYNFPMSSGFRK--KLNRVEBNIDALYWLNDKVF 3533

QY 432 LFKGARVY---VLARGLOVEPYPRSLQD--WGIPEEVSQALPRDGSIIFFRDDRYYR 487

Db 394 LFKSGSGWQMDLARLDIFS---SYKPIKGLFTGVBNPRAAMSGDGRVYFFGKAYYR 450

QY 488 LDOAKLQATTSGRMATLPRMGC 510

Db 451 LNQ-QLRVEKGYPRINSHNMHC 472

## RESULT 2

[illegible]

|                       |                  |  |             |             |
|-----------------------|------------------|--|-------------|-------------|
| Query March           | 18.8%;           | Score 534;   | DB 2;       | Length 669; |
| Best Local Similarity | 27.4%;           | Pred. No. 1.7e-32;   |             |             |
| Matches 167;          | Conservative 64; | Mismatches 206;  | Indels 172; | Gaps 21;    |
| QY                    | 9                | LRAIQLLLMGHLDAPAPERGQELRKAEAEAFLEKYGLINEOVPAKPTSTR-----PSDAI   | 63          |             |
| Db                    | 29               | LLPLLLVLLTGLCTGLGVA--AEDAEVHAENMLRLRYGLTPQ--PSRHNSTYRSQIILASAL | 83          |             |
| QY                    | 64               | RAFOVWSQLVSGLDGDAITLRQMTFRPCGYTDTNYSVAAMERISDLFAHRTYKRRKRP     | 123         |             |
| Db                    | 84               | AEHQRFGLIVTGLDEETKEMMKRPKCGVD-----QFGVRYK-----ANLRBRKRY        | 132         |             |
| QY                    | 124              | AKQGNKYKCHLSYRLVNMPEHLRSRQFAPCAPSPSCGATSQWMSGRPQAPALTSGS       | 183         |             |
| Db                    | 133              | ALTGRKNNHNLFPSTQNYTEKL-----                                    | 155         |             |

```

QY      184 PSSKGTITGM-----AMPLAQAQAPRTPLPR-----      213
Db      156 -----GMYSNEAVRRARFWEQATPLVFOEVPYEDIRLRQKEADIVLPAASGFH      206
QY      214 -----GEAHFDODERMSLSRR--RGRNLFVVLNHEIGTLLG      247
Db      207 GDSPSPDTGTFGLAHAYFPFGPLGGDTHTHDADPEPTFSTDLHGNNLFLVAHVEIGHALIG      266
QY      248 LHSRPRALMLPYKKRLGRDL-LSWDDYLAQSLYGRKG-----GSVAQLPQK      298
Db      267 LEHSSNPNNIAMPFYQWKDVDFKLPBEDLRGICQLYGRPGQOPQTPPLFVTPRRRGR      326
QY      299 LFTDETWDSYSPQ-----GRREPQCKYCHSSFDATV      333
Db      327 P-----DHRPRRPPRRPGKRRPRPKCPVRQPARTEPRDYGNNICDGDPTVAM      379
QY      334 DQOOLYIFKSGHEVAADGNVSE-PRPLQERWVLPRPNIEAAMVSLINDGFYFFKGR      392
Db      380 IR-GMFPYFKGMFWRRVHNRRLVDNYRWPRIGHFMRKGLGDI-SAAVEHQDGRFYFFKGR      437
QY      393 CMRFGRGPRWGLPQ-LCRAG-GLP-RHPDALPPPRRLILTKGARYVYVLARGLQVE      449
Db      438 YLTFEANLBERQYRPLTSYGLGIPYDRIDALIMWEPGHTFFQOEZRYWGFNEETRGD      497
QY      450 PYPYRSLDWMGIDIEVSGALPRPDGS-IIPRDRDYRLDQALQALQATTSRMA TELPYM      508
Db      498 FGYPRIISWGGIPASPKGAFLSNDAAVYTYKGTKYWKFDNELRMBPGYPSKILNDFM      557
QY      509 GCM-HANG      516
Db      558 GCGEHEBEG      566

```

### RESULT 3

matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - rat  
N:Alternate names: membrane-type metalloproteinase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I84471; I61946  
R:Okada, A.; Belluco, J.P.; Rouyer, N.; Chenaud, M.P.; Rio, M.C.; Chambon, P.; Basset, P.  
Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995  
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce  
A:Reference number: I36046; MVID:95224014; PMID:7708715  
A:Accession: I84471  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-582 <RES>  
A:Cross-references: UNIPROT:Q10739; EMBL:X83537; NID:G805012; PIDN:CAA58521.1; PID:G8050  
A:Accession: I61946  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-67, 'M', 69-254, 'A', 256-582 <RE2>  
A:Cross-references: EMBL:X91785; NID:G1001926; PIDN:CAA62897.1; PID:G1001927  
C:Genetics:  
A:Gene: mt-mmp  
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotein  
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-97/Domain: activation peptide #status predicted <PRO>  
F:61-284/Domain: matrix metalloproteinase homology <MMP>  
F:313-508/Domain: hemopexin repeat homology <PN>  
F:93-239, 243, 249/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
F:239, 243, 249/Binding site: zinc, catalytic (His) (active) #status predicted  
F:240/Active site: Glu #status predicted

|    |   |        |                    |        |              |
|----|---|--------|--------------------|--------|--------------|
|    | Query March   | 18.3%; | Score 517.5;       | DB 2;  | Length 582;  |
|    | Best Local Similarity                                       | 29.5%; | Pred. No. 2.56-31; |        |              |
|    | Matches 152; Conservative                                   | 71;    | Mismatches 209;    | Indels | 83; Gaps 20, |
| QY | 37 EAPLEKGY-----LNEGVKPAPTSTFSDAIRAFQVNSQLPVGVDLRATRLRMTPRC | 91     | :                  | :::    | :   ::       |

Db 36 EAMLOOQYGYLPPGDLRTHTORSPOS--LSAALIAIORFYGLQVTKAUSDITKAMRRPRC 93  
 Qy 92 GVTDTNSVAAMARISDLFARHRTQMR---RKRFPAQGNKKYKQHLSTRLVNWPEHLRS 148  
 Db 94 GVPD-----KFTETKANVRKRKRAIQGLKQKHNEITCIGNY----- 131  
 Qy 149 RQFGAPCAPSPSCGATSQR-----WSSGRP-----QPADLTGSSPSSKGTMM---- 192  
 Db 132 -----TPKVGSEYATFEAIRKAFRWESATPLRFREVEYAYIRREGHEKQADIMILFAEG 184  
 Qy 193 --GNAMPMLAQAQAPMTPLPR---RGEAHFPODERWSLSRR--RGRNLFYVLAHEIGT 245  
 Db 165 FHGSDTPPDEGGFLAAHAFPGPNIGDTHFDSAEWTVQNDLNGNDIFLVAVHELGA 244  
 Qy 246 LGLTSPAPRALMAPYKRLG-RDALLSWDVLAVOSLYGKPLGGSVAQLPGKLFDPDR 304  
 Db 245 LGLHNSNDSDIMAFYQMDTENFVLPDDDRGIGQLYGSISGS-----FTKMPQPR 298  
 Qy 305 TWDYSV-POGARPEFQPRYCHSSPDAITVDRQOOLYIFKSHFWEV---AADGNVSEP 359  
 Db 299 TTSRPSVPDKPRNPFTYGPNI CDGNFDTVAMLR-GEWFVFKERWFMVRNNQVMDG---YP 354  
 Qy 360 RPLGRWVGLPPIETAAANSLNDGDFYFKGGRCMRFRPKPVWGLPOLCR--AGLPRH 417  
 Db 355 MFIQGFWRGLPASIN-TAYERKDGKRFVFKGDKHWVFDASDEPGYPKIKELGRGLPTD 413  
 Qy 418 P-DALFPPPLRLILFKGARYVVLARGLQVEPYPRSLQDMGGIPEEVSGLAPRPGS 476  
 Db 414 KIDALLFMWPNCKTYFFRNKITYRFRNEEPRAYVDSIEPKIKVMEGIPESPRGSPGSD 473  
 Qy 477 I-IFPRDRYWRDLQATLQATTSGRWATELPMWGC 510  
 Db 474 FTYFYKGNKYMKFNNOCKLVEPGYPRKSLRDWMC 508  
 Db  
 RESULT 4  
 138028  
 matrix metalloproteinase 14 (EC 3.4.24.-) membrane type precursor - human  
 N/A:alternate names: matrix metalloproteinase MMP14; membrane type matrix metalloproteinase  
 C:Species: Homo sapiens (man)  
 C:Date: 17-May-1996 #sequence revision 17-May-1996 #ext\_change 09-Jul-2004  
 C:Accession: J38028; G02274; J38046; S78011; S45341; S71384  
 R:W11, H.; Hinzmann, B.  
 Eur. J. Biochem. 231, 602-608, 1995  
 A>Title: cDNA sequence and mRNA tissue distribution of a novel human matrix metalloprote  
 A:Reference number: J38028; MUID:95377289; PMID:7649159  
 A:Accession: J38028  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-582 <WILL>  
 A:Cross-references: UNIPROT:P50281; EMBL:248481; NID:G963053; PID:CAA8372.1; PID:G9630  
 R:Luc, G.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: H00963  
 A:Accession: G02274  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-582 <LUO>  
 A:Cross-references: EMBL:U41078; NID:g1127836; PID:AAA83770.1; PID:g1127837  
 R:Okada, A.; Bellio, J.P.; Rouyer, N.; Chénard, M.P.; Rio, M.C.; Chabbon, P.; Basseet, F  
 Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995  
 A>Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce  
 A:Reference number: J38046; MUID:95224014; PMID:7708715  
 A:Accession: J38046  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-7,'S',9-582 <OKA>  
 A:Cross-references: EMBL:X83535; NID:g804993; PID:CAA5519.1; PID:g804994  
 R:Seiki, M.  
 submitted to the EMBL Data Library, January 1994  
 A:Reference number: S78011  
 A:Accession: S78011  
 A:Molecule type: mRNA

A:Residues: 1-7,'S',9-337,'K',339-582 <SEI>  
 A:Cross-references: EMBL:D26512; NID:g793762; PID:BA05519.1; PID:g793763  
 R:Sato, H.; Takino, T.; Okada, Y.; Cao, Y.; Shingawa, A.; Yamamoto, E.; Seiki, M.  
 Nature 370, 61-65, 1994  
 A>Title: A matrix metalloproteinase expressed on the surface of invasive tumour cells.  
 A:Reference number: S45341; MUID:94286011; PMID:8015608  
 A:Accession: S45341  
 A>Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-7,'S',9-188,'R',190,'A',192-267,'K',269-272,'HY',275,'P',277-285,'KQ',288,  
 A:Cross-references: EMBL:D26512  
 R:Sato, H.; Kinoshita, T.; Takino, T.; Nakayama, K.; Seiki, M.  
 FEBS Lett. 393, 101-104, 1996  
 A>Title: Activation of a recombinant membrane type 1-matrix metalloproteinase (MT1-MMP)  
 A:Reference number: S71384; MUID:96397540; PMID:8804434  
 A:Accession: S71384  
 A:Molecule type: protein  
 A:Residues: 112-116 <SNW>  
 C:Genetics:  
 A:Gene: GDB:MMP14; MT1-MMP  
 A:Cross-references: GDB:375731; OMIM:600754  
 A:Map position: 14q11-14q12  
 C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotei  
 C:Keywords: glycoprotein; hydrolase; metalloproteinase; zinc; zymogen  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-97/Domain: activation peptide #status predicted <PRO>  
 F:61-284/Domain: matrix metalloproteinase homology <MMP>  
 F:98-582/Product: matrix metalloproteinase 14 membrane type #status predicted <MAT>  
 F:285-313/Domain: hinge #status predicted <HNG>  
 F:314-508/Domain: hemopexin repeat homology <EPN>  
 F:539-562/Domain: transmembrane #status predicted <TM>  
 F:93,'239,243,249/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:130/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:240/Active site: Glu #status predicted  
 F:319-508/Disulfide bonds: #status predicted

Query Match 18.2%; Score 514.5; DB 2; Length 582;  
 Best Local Similarity 29.7%; Pred. No. 4,26-31;  
 Matches 150; Conservative 69; Mismatches 223; Indels 63; Gaps 18;

Qy 37 EAPLEKYG----LNEOVFKATSTPSDAIRAFQWSQLPVSGVLDAITRQMTPRC 91  
 Db 36 EAMLOOQYGYLPPGDLRTHTORSPOS--LSAALIAIQRFYGLQVTKADADITKAMRRPRC 93  
 Qy 92 GVTDTNSVAAMARISDLFARHRTQMR---RKRFPAQGNKKYKQHLSTRLVNWPEHLRSRQF 151  
 Db 94 GVPD-----KFTETKANVRKRKRAIQGLKQKHNEITCIGNY----- 135  
 Qy 152 GAPCAPSPSCGATSQWSSGRP-----QPADLTGSSPSSKGTMM-----GNAMPMA 200  
 Db 136 -GEVATFEALIRKAFVWESATPLRFREVEYAYIRREGHEKQADIMLFAEGFHDGDFRFG 194  
 Qy 201 OGAPMTPLPR---RGEAHFPODERWSLSRR--RGRNLFYVLAHEIGTGLTSPAPR 255  
 Db 195 ECGFLAAHYFPGPNIGDTHFDSAEWTVQNDLNGNDIFLVAVHELGAQLHESSSDS 254  
 Qy 256 ALMAPYKRLG-RDALLSWDVLAVOSLYGKPLGGSVAQLPGKLFDPDR 313  
 Db 255 ALMAPYQMDTENFVLPDDDRGIGQLYGSISGS-----FTKMPQPRRTSRSPVDPK 308  
 Qy 314 RRPETQPRYCHSSPDAITVDRQOOLYIFKSHFWEV---AADGNVSEPRLOERWGL 369  
 Db 309 PKNPYTPGPNICDGNFDTVAMLR-GEWFVFKERWFMVRNNQVMDG---YPMFIQGFWRGL 364  
 Qy 370 PNIEAANSLNDGDFYFKGGRCMRFRPKPVWGLPOLCR--AGLPRH-PDALFPP 426  
 Db 365 PASIN-TAYERKDGKRFVFKGDKHWVFDASDEPGYPKIKELGRGLPTD 423  
 Qy 427 LARLLFKGARYVVLARGLQVEPYPRSLQDMGGIPEEVSGLAPRPGSI-IFFRDRY 485  
 Db 424 NGKTYFFRNKITYRFRNEEPRAYVDSIEPKIKVMEGIPESPRGSPGSD 483



A/Content: Skin wounds  
 A/Accession: J06197  
 A/Molecule type: mRNA  
 A/Residues: 1-491 <OKA>  
 A/Cross-references: UNIPROT:P97568; GB:U46034  
 C/Comment: This protein is a member of the matrix metalloproteinase family.  
 C/Superfamily: Interstitial collagenase; homopeptin repeat homology; matrix metalloprotease  
 C/Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F/52-261/Domain: matrix metalloproteinase homology <PXX>  
 F/291-483/Domain: homopeptin repeat homology <PXX>  
 F/84-218, 222, 228/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F/219, 222, 228/Binding site: zinc, catalytic (His) (active) #status predicted  
 F/219/Active site: Glu #status predicted

Query Match 16.3%; Score 463; DB 2; Length 491;  
 Best Local Similarity 29.2%; Pred. No. 2, 7e-27;  
 Matches 153; Conservative 59; Mismatches 232; Indels 80; Gaps 22;

QY 2 VARVGLLRA-----LQLLMGHL--DAQPARGGQELRKAFLKRGY 45  
 DB 1 MAAPALRLAIRALLPLRLILLPLRLPMADRPENRHRPRVKVPLL--AA 58  
 QY 46 LNEQVFKAPTSTRFSADAIRAFQWVSQLPVSGVLDRAITLRQMTPE--RCGVTDITNSYAMAE 104  
 DB 59 LPLSLISVPS-----HNVPG--PASS-----SRPLRGVDPDDV----- 92  
 QY 105 RISDLFARHRTKMRKKRPAKQGNKYKQHLSTRLVNPENHLSRQGFACAPSSCGAT 164  
 DB 93 -----LNARN-----RQKRFVLSGGRMEKTDLTTRIILFPWQLVREQVQTVA-----EA 137  
 QY 165 SQWSSSGRPPQAPRLTSGSSSKGTTMGW---AMPMAQAPWRTPFL---RRGAHF 218  
 DB 138 LRAWSEVTLPTFTVEHGRADIMIDFRVWNGDLFPDGGGILANAFFKTHREGDVHF 197  
 QY 219 DQDERSLSRRRGNLFVLAHEIGTGLTSPAPRALMAPYKRLGRDALISWDVLA 278  
 DB 198 DYBETWTIG--DKGTDLQVAHHEGNYLGIQHTTAALMSPFY--TRYPILSLSPDDRGG 255  
 QY 279 VQSLYKPKLQGSVAVALPGKLTDFETWDSYSPQRRPETQPKYCHSSFDATITVDROQ 338  
 DB 256 IQLHYGR---QLPTSPPTLSQAQGTDTNEIALQPEEVP--PEVCETSPDAVSTIR--GE 310  
 QY 339 LYLFGKSHFWEVADGNVSE--PRPDERVNGLPENIEAANAISLNDQDFPFKGRGCMRF 396  
 DB 311 LPEFKAGFVWRLES--GQLQGYRALASRHQGLSPVDALFEDE--QGIWFFQAGQWVY 368  
 QY 397 RGPFPVWGLPOLCRAG--GLPRHPDALFPPLRLILFKGARYVVLARGLAQVEPYPR 454  
 DB 369 DGEKFPVLPRLSLGLQGSVVAH--ALVWGEKKIKITFFRGDVTWRHPRTRQVNDNVPR 427  
 QY 455 SLQDMGIPPEVSGALPRPDGSIIFFRDRIYRLDQAKLTATTS 498  
 DB 428 RTTDMRGVPSIDIAFQDAEGYAYFLRGHLWKFPVKVLTES 471

RESULT 8  
 stromelysin 3 (EC 3.4.24.-) precursor - human  
 N/Alternate names: matrix metalloproteinase 11 (MMP11)  
 C/Species: Homo sapiens (man)  
 C/Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
 C/Accession: S13423; I38250; S58912  
 R/Basset, P.; Belloq, J.P.; Wolf, C.; Stoll, I.; Huch, P.; Limacher, J.M.; Podhajcer, Nature 348, 699-704, 1990  
 A/Title: A novel metalloproteinase gene specifically expressed in stromal cells of breast  
 A/Reference number: S13423; MUID:91080920; PMID:1701851  
 A/Accession: S13423  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-488 <BAS>  
 A/Cross-references: UNIPROT:P24347; EMBL:X57766  
 R/Argland, P.; Meiot, T.; Guerin, E.; Thomas, G.; Basset, P. J. Biol. Chem. 270, 20337-20344, 1995

A/Title: Structure and promoter characterization of the human stromelysin-3 gene.  
 A/Reference number: I38250; MUID:95386471; PMID:7657606  
 A/Accession: I38250  
 A/Status: preliminary; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-36 <RBS>  
 A/Cross-references: EMBL:X84664; NID:g984746; PIDN:CAA59150.1; PID:g987949  
 R/Pei, D.; Weiss, S.J. Nature 375, 244-247, 1995  
 A/Title: Furin-dependent intracellular activation of the human stromelysin-3 zymogen.  
 A/Reference number: S58912; MUID:95265105; PMID:7746527  
 A/Accession: S58912  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 81-101 <PEI>  
 C/Genetics:  
 A/Genes: GDB:MMP11; STWY3  
 A/Cross-references: GDB:128630; OMIM:185261  
 A/Map position: 22q11.2-22q11.2  
 C/Superfamily: Interstitial collagenase; homopeptin repeat homology; matrix metalloprotease  
 C/Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F/48-258/Domain: matrix metalloproteinase homology <PXX>  
 F/488-480/Domain: homopeptin repeat homology <PXX>  
 F/80, 215, 219, 225/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F/215, 219, 225/Binding site: zinc, catalytic (His) (active) #status predicted  
 F/216/Active site: Glu #status predicted

Query Match 16.1%; Score 455.5; DB 2; Length 488;  
 Best Local Similarity 30.0%; Pred. No. 9, 8e-27;  
 Matches 136; Conservative 51; Mismatches 202; Indels 65; Gaps 17;

QY 68 WVSQLPVSGVLDRAAT-----LRQMTPECGVTDITNSYAMABRISDLFARHRTKMRKK 121  
 DB 51 WNAALSSPARATQEARPPASLAPRPGVDPD-----DGLSARN-----RQK 96  
 QY 122 RPAKQGNKYKQHLSTRLVNPENHLSRQGFACAPSSCGATSQWSSGRPPQAPLTS 181  
 DB 97 RFLSGGRMEKTDLTTRIILFPWQLVREQVQTVA-----EALKWSDVTPLTFTVEH 150  
 QY 182 GSPSSKGTITMGW---AMPMAQAPWRTPFL---RRGAHFDQDERWSLSRRGRNLF 235  
 DB 151 GRADIMIDFARYWDGDLFPDGGGILANAFFKTHREGDVHFDYDTWTIGDQGTDL 210  
 QY 236 VVLAHEIGTGLTSPAPRALMAPYKRLGRDALISWDVLAVALQSLYGR----- 286  
 DB 211 QVAHHEGNYLGIQHTTAALMSAFY--TRYPILSLSPDCCRGVGHLYGQRPWTVTSRTP 269  
 QY 287 -LGSAVAVQLPGKLTDFETWDSYSPQRRPETQPKYCHSSFDATITVDROQLYIFKGS 345  
 DB 270 ALGPQAGI-----DTNEIALPEPDAF--PDACASPDVAVSTIR--GLIFPFKAG 314  
 QY 346 HFWEVADGNVSE--PRPDERVNGLPENIEAANAISLNDQDFPFKGRGCMRFPRPVW 403  
 DB 315 FVWRLRG--GQLQGYRALASRHQGLSPVDALFEDE--QGIWFFQAGQWVVDGKFPVL 372  
 QY 404 GLPOLCRAGGLPRHP--DAALFF--PPLRLILFKGARYVVLARGLAQVEPYPSLQDMG 461  
 DB 373 GPAPLTLL--GLVFPVVAALVWGEKKIKITFFRGDVTWRHPRTRVDSVPRRATWVRG 431  
 QY 462 IPREVSGALPRPDGSIIFFRDRIYRLDQAKLQA 495  
 DB 432 VPSEIDIAFQDADQYAYFLRGRLYWKFDPVYKVA 465

RESULT 9  
 KCMSS1  
 stromelysin 1 (EC 3.4.24.17) precursor - mouse  
 N/Alternate names: collagenase activating protein; matrix metalloproteinase 3 (MMP3); p  
 C/Species: Mus musculus (house mouse)  
 C/Date: 30-Sep-1992 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C/Accession: J01476; S18867; B32963; S33139  
 R/Hammani, K.; Henriet, P.; Beckhout, Y. Gene 120, 321-322, 1992

[illegible]

Db 309 - FKDRHPM- ----KSLRTDEPEFYLLISFWPSLPNNDAAYEVYTRNDTVFIKNGQWAI 363

QY 337 RQPKPVGGLPOLCRAGGLP---RHPPALFFPPLRLILFKGARYYLLANGGLQVEEYP 453

Db 364 RGHHEIAGPYKSIHTLGLPATVTKIDDAISNKKRTYFFVEDKYMRFDEKOSMEGFP 423

QY 454 RSL-ODMGSIPEEYVGLPRPDGSIILFFPRDRYRWLDQAKLTQTTSGRMATEL-----PMM 508

Db 424 RKIADDFGVDSRDV-AVFEARGFLYFFS-----GSSQLPEPDPAKXKTYTHLKSNSWF 475

QY 509 GC 510

Db 476 NC 477

RESULT 10

KCRITH

stromelysin 1 (EC 3.4.24.17) precursor - rat

N/Alternate names: collagenase activating protein; matrix metalloproteinase 3 (MMP3); p

C/Species: Rattus norvegicus (Norway rat)

C/Date: 13-Aug-1986 #sequence revision 13-Aug-1986 #text\_change 09-Jul-2004

C/Accession: A00997; P50150; S22767

R/Matrisian, L.M.; Glatthorn, N.; Gesnel, M.C.; Breathnach, R.

EMBO J. 4, 1435-1440, 1985

A/Title: Epidermal growth factor and oncogenes induce transcription of the same cellula

A/Reference number: A00997; MUID:85284930; PMID:3875462

A/Accession: A00997

A/Molecule type: mRNA

A/Residues: 1-475 <MA1>

A/Cross-references: UNIPROT:P03957; GB:X02601; NID:957460; PIDN:CAA26448.1; PID:957461

J. Umenishi, F.; Yasumitsu, H.; Ashida, Y.; Yamauti, J.; Umeda, M.; Miyazaki, K.

R. Biochem. 108, 537-543, 1990

A/Title: Purification and properties of extracellular matrix-degrading metallo-proteina

A/Reference number: P50150; MUID:91154156; PMID:1963430

A/Accession: P50150

A/Molecule type: protein

A/Residues: 19-20, 'X', 22-28, 110-112, 'X', 114-115, 'X', 117, 'X', 119, 309-325 <UME>

R. Breathnach, R.; Matrisian, L.M.; Gesnel, M.C.; Staub, A.; Leroy, P.

Nucleic Acids Res. 15, 1139-1151, 1987

A/Title: Sequences coding for part of oncogene-induced transin are highly conserved in a

A/Reference number: A26403; MUID:87146421; PMID:3547333

A/Contents: annotation; introns

A/Note: Intron positions were determined by comparison of the previously reported cDNA s

R. Sanchez-Lopez, R.; Nicholson, R.; Gesnel, M.C.; Matrisian, L.M.; Breathnach, R.

J. Biol. Chem. 265, 11892-11899, 1988

A/Title: Structure-function relationships in the collagenase family member transin.

A/Reference number: S22767; MUID:88298869; PMID:2841336

A/Contents: annotation; active site; activation

A/Note: molecules with mutations in the autocatalytic region showed a much increased te

R. Park, A.J.; Matrisian, L.M.; Kells, A.F.; Pearson, R.; Yuan, Z.; Navre, M.

J. Biol. Chem. 265, 1584-1590, 1991

A/Title: Mutational analysis of the transin (rat stromelysin) autocatalytic region demon

A/Reference number: A43028; MUID:91107652; PMID:1988438

A/Contents: annotation; autocatalytic region

A/Note: Arg-89 and Cys-92 are essential for maintaining latency

C/Comment: This enzyme degrades various extracellular matrix proteins, including fibronect

C/Comment: Stromelysin 1 hydrolyzes peptide bonds in plasminogen to yield a fragment with

C/Comment: Stromelysin 1 activates its proenzyme after cleavage(s) within the activation

C/Comment: Prostromelysin is found in glycosylated and unglycosylated forms, both of whic

C/Genetics:

A/Intron: 33/3; 115/2; 165/1; 207/1; 262/1; 310/2; 355/1; 408/2; 443/1

C/Function:

A/Description: endopeptidase preferentially hydrolyzing peptide bonds on the carboxyl side

C/Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprote

C/Keywords: calcium; extracellular matrix; fibronectin; glycoprotein; hydrolase; metallo

F.1-17/Domain: signal sequence #status predicted <SIG>

F.18-475/Product: prostromelysin 1 #status predicted <PRO>

F.18-97/Domain: activation peptide #status predicted <ACT>

F.158-262/Domain: matrix metalloproteinase homology <MMP>

F.188-95/Region: matrix metalloproteinase homology <MMP>

F.98-475/Product: stromelysin 1 #status predicted <MAT>

F.282-475/Domain: hemopexin repeat homology <PXN>

F:90,216,220,226/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:118/Binding site: carboxypeptidase (Asn) (covalent) #status predicted  
 F:216,220,226/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:217/Active site: Glu #status experimental  
 F:288-475/Diulfide bonds: #status predicted

Query Match 15.1%; Score 427; DB 1; Length 475;

Best Local Similarity 28.1%; Pred. No. 1,4e-24;  
 Matches 142; Conservative 69; Mismatches 203; Indels 92; Gaps 20;

```

QY 9 LRALQLLLW-----GHLDAPAEKGGDELRKEAEFLKRYGYLNEQVPAKSTRRS 60
D 1 MKGLPVLMLCTAVCSSYPLHSEEDAGMEVLQK-----YLENYGLKDVYQFKKKDSS 56
QY 61 DALRPFVWGO---LPVSGVLDRATLQMTFRPGVDTNYSYAMARISDLFARHRTM 117
D 57 PVVKKIEMOKFLGLKMTGKLDSNTMELMKPRGVDVGGSTF----- 101
QY 118 RRRKRFPAQGNKMYKQHLRYLVNMPHLSRQFGAPCAPSSCGATSGR-----WSSGRP 173
D 102 -----PGSPKMKKMHSTRIVYITDL-----PRESVDSALERALKWMEVTP 144
QY 174 QPAPLITSGSPS---SKGTTMGWAMP-LMAQAPWRTFPLP---RGEAHPDODERWLS 227
D 145 LTFGRISGEADIMISPAVEEHGDFIPDGGMTLAHAAYAGPGTNGDAHPDDDERWT-D 203
QY 228 RRRRRNLPVLAHEIGTTLGTHSPAPRALMAPYTKR---LGRDALLSMDVLAVOSLY 284
D 204 DVTGTNPLVLAHHEIGSLGPHSNAMEALMYPYKSTDLAR-FHLSQDDVDGIQSLYG 262
QY 285 KPLGGSVAVOLPGKLTDFETWDSYSPQGRPETQPKYCHS--SPDAITVDRQOOLYI 342
D 263 PPTSSPDVLPVPTK-----SNSLDPELT--PMCSALSFDVASTLRGEVLP-F 307
QY 343 KGSHEWEVADGNVSEBRP---LQERWVGLPNIIEAAVSLNDGDFYFKGRCMFRPG 398
D 308 KDRFPMR---KSLRTPRPFYLLSSFWPSLPNMMDAAYEVYTRNDYVILKGNQIMAIRG 363
QY 399 PKPFWGLPQLCRAGLP---RHPPALPFPPLRLILFKARYVYVLARGLQVEPYPRS 455
D 364 HEELAGYPKSITHLGLPETYKIDALISLKDKQKTYFVEDKFWRFDEKKSMDPEFRK 423
QY 456 L-ODMGIPREVSGLPRPDGSIIF 480
D 424 IAEHPGIGTKVD-AVFEAFGLYIF 448

```

# RESULT 11

A53711  
 collagenase 3 (EC 3.4.24.-) - human  
 N/Alternate names: matrix metalloproteinase 13 (MMP13)  
 C/Species: Homo sapiens (man)  
 C/Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004  
 C/Accession: A53711  
 R/Breathnach, J.M.P.; Dierckx, I.; Balbin, M.; Sanchez, L.M.; Blasco, R.; Tolliva, J.; Lof, J. Biol. Chem. 269, 16766-16773, 1994  
 A/Title: Molecular cloning and expression of collagenase-3, a novel human matrix metalloproteinase  
 A/Reference number: A53711, MUID:94266894, PMID:8207000  
 A/Accession: A53711  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-471 <FRS>  
 A/Cross-references: UNIPROT:P45452; GB:X75308; NID:9516385; PIDD:CAA53056.1; PID:9516386  
 C/Genetic8:  
 A/Genes: GDB:MMP13; CLG3  
 A/Cross-references: GDB:373966; OMIM:600108  
 A/Map position: 11q22.2-11q22.3  
 C/Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloprotei  
 C/Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F:64-471/Domain: matrix metalloproteinase homology <MMP>  
 F:274-471/Domain: hemopexin repeat homology <HXN>  
 F:96,222,226,232/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:222,226,232/Binding site: zinc, catalytic (His) (active) #status predicted

F:223/Active site: Glu #status predicted

Query Match 15.0%; Score 426.5; DB 2; Length 471;

Best Local Similarity 30.0%; Pred. No. 1.5e-24;  
 Matches 151; Conservative 55; Mismatches 204; Indels 93; Gaps 19;

```

QY 8 LRLALQLLLWGHLDAPAEKGGDE--LRKEAEFLKRY-----GYLNEQVPAK 54
D 5 VLAAPFLSWTHCRALPLPSGDEDDLSIEDLQFAEYILSYHPTNLADILKENAASM 64
QY 55 TSTRFSALTAFPWQSLPVSGLDRAATLQMTFRPGVDTNYSYAMARISDLFARH 114
D 65 T-----ERLEKMSPPGLEVTGKLDDVTLDVMKKPRGVDVGBRYVFPRTL----- 111
QY 115 TKRRKRKRFPAQGNKMYKQHLRYLVNMPHLSRQFGAPCAPSSCGATSGRWSGRP 173
D 112 -----KSKMNLTYRIVYITDMTHS-----VEKAFKAAPKWSVDTP 150
QY 174 QPAPLITSGSPS---SKGTTMGWAMP---LMAQAPWRTFPLRGEAHPDODERW 224
D 151 INFTRLHDGIADIMISFGIKRHGDYFPDGPGLAHAFP---FGPNYGGDAHPDDDET 207
QY 225 SLRRRRNLPVLAHEIGTTLGTHSPAPRALMAPYTKRLGRD-ALLSMDVLAVOSLY 283
D 208 T-SSSKGYNPLVLAHHEIGSLGDSKDGALMFPITYTYSKSHFMLPDDVQIGSLY 266
QY 284 GKPLGGSVAVOLPGKLTDFETWDSYSPQGRPETQPKYCHS--SPDAITVDRQOOLYI 341
D 267 G-----PG-----DEPNRKHPT--PDKCDPSLSIDALITSLRGERM-I 302
QY 342 KGSHEWEVADGNVSEBRPLQERWVGLPNIIEAAVSLNDGDFYFKGRCMFRGPKP 401
D 303 FKORFPRHLHPQVDALFLTKSPWPLPRIDAAVSHPSHDLIFIRGKRFMALNGDI 362
QY 402 VMGLPQLCRAGLP---DAALFPPPLRLILFKARYVYVLARGLQVEPYPRS-LQ 457
D 363 LEQYPKISBELGPKVEYKISAAVHFEDTGKTLFSGNQVWRVDTNTHMDKDPRLIE 422
QY 458 DMGIPREVSGLPRPDGSIIF 480
D 423 DPEIGDKVD-AVYERKGYIIF 444

```

# RESULT 12

KCRS2  
 stromelysin 2 (EC 3.4.24.22) precursor - rat  
 N/Alternate names: matrix metalloproteinase 10 (MMP10); transin-2  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004  
 C/Accession: B26403; A41775; S26498  
 R/Breathnach, R.; Matrisian, L.M.; Gesnel, M.C.; Steub, A.; Leroy, P.  
 Nucleic Acids Res. 15, 1139-1151, 1987  
 A/Title: Sequences coding for part of oncogene-induced transin are highly conserved in  
 A/Reference number: A26403; MUID:87146421; PMID:3547333  
 A/Accession: B26403  
 A/Molecule type: mRNA  
 A/Residues: 1-476 <BBR>  
 A/Cross-references: UNIPROT:P07152; EMBL:X05083; NID:957388; PIDD:CAA28739.1; PID:95738  
 A/Note: Inton positions were determined by comparison of the cDNA sequence to genomic  
 A/Note: mRNA for this protein was expressed in several transformed rat embryo fibroblasts  
 R/Chan, J.C.; Scanlon, M.; Zhang, H.Z.; Jia, L.B.; Yu, D.H.; Hung, M.C.; French, M.; Ba  
 J. Biol. Chem. 267, 1099-1103, 1992  
 A/Title: Molecular cloning and characterization of v-mos-activated transformation-asso  
 A/Reference number: A41775; MUID:92112740; PMID:1370458  
 A/Accession: A41775  
 A/Molecule type: mRNA  
 A/Residues: 1-476 <CHA>  
 A/Cross-references: GB:M65253; NID:9207150; PIDD:AAA42202.1; PID:9207151  
 A/Note: Sequence extracted from NCBI backbone (NCBI:76184)  
 R/De Vong, M.W.; Mukherjee, B.B.  
 Oncogene 7, 109-119, 1992  
 A/Title: Transformation of normal rat kidney cells by v-K-ras enhances expression of tr  
 A/Reference number: S26496; MUID:92158347; PMID:1741158



A/Accession: S26498  
A/Status: preliminary; translation not shown  
A/Molecule type: mRNA  
A/Residues: 31-103, 'L', 241-242, 'TOMEKPR', 251, 'L', 253-254, 'CE', 293-294, 'L', 296 <DEV>  
A/Cross-references: EMBL:X64020  
C/Genetics:  
A/Intons: 35/3; 117/2; 167/1; 209/1; 264/1; 311/2; 356/1; 409/2; 444/1  
C/Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotease  
C/Keywords: calcium; extracellular matrix; fibroblast; glycoprotein; hydrolase; metalloprotease  
F/1-17/Domain: signal sequence #status predicted <SIG>  
F/18-99/Domain: prostromelysin 2 #status predicted <PRO>  
F/60-264/Domain: activation peptide #status predicted <ACT>  
F/90-97/Region: autolysosomal homology <AMP>  
F/100-476/Domain: hemopexin repeat homology <MAT>  
F/120-218, 222, 228/Binding site: zinc, catalytic (His, His, His) (inhibited) #status predicted  
F/218, 222, 228/Binding site: zinc, catalytic (His) (active) #status predicted  
F/219/Active site: Glu #status predicted  
F/289-476/Disulfide bonds: #status predicted

Query Match  
Best Local Similarity 15.0%; Score 425.5; DB 1; Length 476;  
Matches 152; Conservative 59; Mismatches 237; Indels 79; Gaps 19;

QY 9 LRAIQLLMGLDQAP---AERGOELKKEAEAELEKYGYL---NEQVPKAPTSTRSDA 62  
DB 4 LAIVLVCFLPCISVPLHGAVRDHSHTDLAQVLEKYNFRKNEKPFKRXSSPVVK 63  
QY 63 IRAQVWSQPLVSGVLDRLATLRQMTPRCGVTDNSVAAMERISDLFAHRTMKRKR 122  
DB 64 IEEMQKFLGLEMTGKLDNSNTVMHMKRCVDPVGGSTP----- 103  
QY 123 FAKGKNKMYQKHSYLRVMPPEHLRSKQFAPCAPSSCGATSOR---WSSGRPOQAP 178  
DB 104 --PQSPKRNKHSYLRVNTYLDL-----PRESVSAIERLAKWEETPLTPSR 151  
QY 179 LITGSPBS---SKGTTTGMAMPPLMAOGAPWRTPLR---RGAHFDODERMSLRNR 232  
DB 152 ISEGDADIMISFVAGHEGDFPFDVGOSLAHAYPPGPGYGDHDDDEKSLG--PQGT 210  
QY 233 NLFVYLAHEIGTGLTSPAPRALMAPYKRLGRDA--LISWDVLAQSLYLG--KPLCG 289  
DB 211 NLFVYLAHEIGTGLTSPAPRALMAPYKRLGRDA--LISWDVLAQSLYLG--KPLCG 289  
QY 290 SVAVOLPGKLTDFEWDYSPOGRREPTQGRYCHS--SPDAITVDROOQLYFKGSHF 347  
DB 271 AIVVPEP-----SVSP---KPER--PVKCDPALSFDAVTLNGEPLF--PKDHP 313  
QY 348 WEVAADGNVSEPRPLQERWGLPPIIEAASVSLNCGDFYFGGRCMFRGKPYWGLPQ 407  
DB 314 WRTQWNPBEPFHLISAFWPSLPSGLDAAYANNDKRVLFKSGQFMAVRNGEVQAGYPK 373  
QY 408 LCRAGLP---RHPDALFPPPLRLILFKGARVYVLAAGLOVEPPYRSL--QDWGIR 463  
DB 374 RIHTLGFPPYVKKIDAAVFEKKEKKTYFFVGDKXWRDETRQMDKGFRLITDPFGIE 433  
QY 464 BEVSGALPRPDGSIIFRDDRVLDDQAKLQATTSGRWATELPMWGC 510  
DB 434 PQVDAYL-HARGFYFFGSSQGFDP--NARTVHTLXNSWILC 476

## RESULT 13

JC5743  
matrix metalloproteinase (EC 3.4.24.-) precursor - pig  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 09-Dec-1997 #sequence\_revision 23-Jan-1998 #text\_change 09-Jul-2004  
R/Accession: JC5743  
R/Bartlett, J.D.; Simmer, J.P.; Xue, J.; Margolis, H.C.; Moreno, E.C.  
Gene 183, 123-128, 1996  
A/Title: Molecular cloning and mRNA tissue distribution of a novel matrix metalloproteinase  
A/Reference number: JC5743, MUID:97149288; PMID:8996096

A/Accession: JC5743  
A/Molecule type: mRNA  
A/Residues: 1-483 <BAR>  
A/Cross-references: UNIPROT:P79287; GB:U54825; NID:g1800212; PIDN:AB41396.1; PID:g1800  
C/Comment: This enzyme plays a role in enamel biomineralization and development.  
C/Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotease  
C/Keywords: hydrolase; metalloproteinase; zinc  
F/1-22/Domain: signal sequence #status predicted <SIG>  
F/23-483/Domain: matrix metalloproteinase #status predicted <MAT>  
F/68-271/Domain: matrix metalloproteinase homology <AMP>  
F/100-226, 230, 236/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status predicted  
F/226, 230, 236/Binding site: zinc, catalytic (His) #status predicted  
F/227/Active site: Glu #status predicted

Query Match  
Best Local Similarity 14.5%; Score 410; DB 2; Length 483;  
Matches 143; Conservative 64; Mismatches 209; Indels 86; Gaps 21;

QY 36 AEAFLEKY-----GY-LNEQVPKAPTSTRSDAIRAFQWSQPLVSGVLDRLATLRQMT 89  
DB 41 AQAYLQKTYTKKGGHGVGEVAKGNS--MVKKIKELQAFFGIRVTKLDRITMDVTKR 98  
QY 90 RCGVTDNSYAAWAEIRISDLFAHRTMKRKRFAKQKNKYQKHSYLRVMPPEHLRSR 149  
DB 99 RCGVDPVANY-----RLPFGPKMKKNTLYTISKYTSM--- 133  
QY 150 QFGAPCAPSSCGATSORWSSGRPOQAPLITGSPS---SKGTTTGMAMPPLMAOGAPWR 206  
DB 134 ---TAEVDKAMEMALQAWSSAVPLSPVRVNAGEADIMISFETGDHGSYPDPGRGLTA 190  
QY 207 TFFLPRR---GEAHFDODERMSLRNRANLFFVLAHEIGTGLTSPAPRALMAPYK 263  
DB 191 NAFAPGEGIGDPTTHDNEKMTWG--MGPNLFTVAHHEFGALGASHDTSALMVPYK 249  
QY 264 RLGRDAL--LSMDVLAVALSLYGLKPLGSGVAVOLPGKLTDFEWDYSPOGRREPTQGR 322  
DB 250 YQNPYGFHLPRDQVKGIALYV-----PRKFTGKPT-----VPHRPNNPSLPD 294  
QY 323 YC--HSSPDATTVNROOQLYFKGSHF--EVAADGNVSEPRPLQERWGLPPIIEAASV 378  
DB 295 ICDSSSGSDATVM--LGKELFFRDRITWRQVHLSGI--PSTITSSFPQLMSVDAAYE 352  
QY 379 SLNDGDFYFFKGGRCMFRGKPYWGLPOLCRAGLGRHP--DAALFPPPLRLILFKG 435  
DB 353 VADRGMAVYFFKGGRYWTRGPO--MGSPRITVDFGFFRYQRIIDAAVHLKDTQKTLFVG 411  
QY 436 ARVYVLAAGLOVEPPYRSL--QDWGIRPEVSGALPRPDGSIIFRDDRVLDDQAK-- 492  
DB 412 DEITYSDERKRKNDKQYPKNTEEFSGVNGQIDAAY--ELNGYIYFSGPRAKYDTREKED 470  
QY 493 ----LQATTSGRWATELPMWGC 510  
DB 471 VVSLKNS-----WICG 483

## RESULT 14

A23685  
interstitial collagenase (EC 3.4.24.7) precursor - rat (fragment)  
N/Alternate names: matrix metalloproteinase 1 (MMP1); vertebrate collagenase  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 09-Jul-2004  
R/Quinn, C.O.; Scott, D.K.; Brinckerhoff, C.E.; Mattislian, L.M.; Jeffrey, J.J.; Partridge  
J. Biol. Chem. 265, 22342-22347, 1990  
A/Title: Rat collagenase. Cloning, amino acid sequence comparison, and parathyroid hormone  
A/Reference number: A23685; MUID:91093077; PMID:2176215  
A/Accession: A23685  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-466 <QUT>  
A/Cross-references: UNIPROT:P23097; GB:M60616; GB:M36452; NID:g203498; PIDN:AA72124.1; I



C:Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloproteinase; hydrolase; metalloproteinase; zinc; zymogen  
 C:Keyword: hydrolase; metalloproteinase; zinc; zymogen  
 F:159-262/Domain: matrix metalloproteinase homology <MP>  
 F:1273-466/Domain: hemopexin repeat homology <PXN>  
 F:91,217,221,222/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:217,221,222/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:218/Active site: Glu #status predicted

Query Match 14.4%; Score 407.5; DB 2; Length 466;  
 Best Local Similarity 28.1%; Pred. No. 3,9e-23;  
 Matches 147; Conservative 65; Mismatches 214; Indels 97; Gaps 20;

```

14 LLLMGLDADQAPRG---GQELRKE---AEAFLEKY-----GYLNDQVPKAPSTSRF 59
5 LLSWTHCWSLPLRYGDDDDDLSEEDLEFAEHYLSKYHVPVTLAAIL-----KKSIVTST 59
60 SDAIRAFQVNSQLPVSGVLDRAITLRQMTPRCGVTTNSYAAAEKISDLFARHRTKMR 119
60 VRLRERMQSPFGIDVYTKLDDPTLDIRKPRCGVPVGYVNPFRTL----- 106
120 KRRFAKQGNKMYKHLSTRLVNWPEHLRSQFGAPCAPSSCGATSGRMSGRPOQAPL 179
107 -----KMSQTNITVRIVNTPTDISSEV-----EKAPRKAKKMSDVTPPLNFTRI 151
180 TSGSPS---SKGTTTGMAMP-----LMAQGAPEWTPPLPRGGAHFDODERMSLSRRR 230
152 HDGTADIMISFGTKEHGDFFPDGSGLLAHAPF---PGNLCGDAHFDDEFTWT-SSSK 207
231 GNNLPLVLAHEHGTITGLTHSPAPRALMAPYTKLRD-ALLSMDVLAVALSYKPLGG 289
208 GNNLPLVLAHEHGTITGLTHSPAPRALMAPYTKLRD-ALLSMDVLAVALSYKPLGG 262
290 SVAVALPGKLTFTDFETWDSYFQGRAPETQGRPKYCHS--SPDAITVDRQOQLYFKGSHF 347
263 -----PG-----DEDPNPKPKPT--PEKCDPALSLDAITLSRGETM-IFKDRPF 303
348 WEVAADGNVSEPRPLGERWVGLRPNI EAALVSLNDGDFYFKGRCMRFRGPPVNLQ 407
304 MRLHPQOVPELPLTSGFWPELPHNVDAVAEHPSRLMFLPRGRKFMALNGVYIMGYPR 363
408 LCRAGGLPRHP--DAALFPPLRLILKRGARYTLARGGLQVETRYRSL-QDMGGIP 463
364 KISDLFPKRVKRLSAVAHFEDTGKTLFPSGNHWMSYDANDMDXDYRLIEEPPGIG 423
464 EEVSAGLPRPDGSIIFPR-----DRIYMLDQAKLQATTSGRW 501
424 DKVD-AVYERKGIYTFNCPIDPFISIMSNRIYRVNPTNSLW 465

```

## RESULT 15

JC6505  
 Stromelysin 2 (EC 3.4.24.22) precursor - mouse  
 M:Alternate names: matrix metalloproteinase 10  
 C:Species: Mus musculus (house mouse)  
 C>Date: 16-Oct-1998 #sequence\_rev1510n 16-Oct-1998 #text\_change 09-Jul-2004  
 C:Accession: J06505  
 R:Madlener, M.; Werner, S.  
 Gene 202, 75-81, 1997.  
 A>Title: cDNA cloning and expression of the gene encoding murine Stromelysin-2 (MMP-10).  
 A:Reference number: J06505; M0ID:98087420; PMID:9427548  
 A:Accession: J06505  
 A:Molecule type: mRNA  
 A:Residues: 1-476 <MAD>  
 C:Cross-references: UNIPROT:O55123; GB:Y13185; NID:g2793311; PIDN:CAA73641.1; PID:g27913  
 C:Comment: This enzyme degrades various extracellular matrix proteins, including fibronectin.  
 C:Gene: MMP-10  
 C:Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloproteinase; hydrolase; metalloproteinase; zinc; zymogen  
 C:Keyword: hydrolase; metalloproteinase; zinc; zymogen  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-476/Product: Stromelysin 2 #status predicted <PRO>  
 F:18-99/Domain: activation peptide #status predicted <ACT>  
 F:60-264/Domain: matrix metalloproteinase homology <MP>

F:90-97/Region: autoinhibitory  
 F:100-476/Product: Stromelysin 2 #status predicted <MAT>  
 F:283-476/Domain: hemopexin repeat homology <PXN>  
 F:92,218,222,228/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:120/Binding site: carbonylate (Asn) (covalent) #status predicted  
 F:218,222,228/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:219/Active site: Glu #status predicted  
 F:289-476/Dileuflide bonds: #status predicted

Query Match 14.3%; Score 406; DB 1; Length 476;  
 Best Local Similarity 27.5%; Pred. No. 5.3e-23;  
 Matches 138; Conservative 68; Mismatches 211; Indels 84; Gaps 19;

```

36 AEAFLEKYGL---NEQVKAFTSTFSDAIRAFQVNSQLPVSGVLDRAITLRQMTPRCG 92
34 AQGYLEKYVNFKNKQKIFKRDSPPVYKIDEMQKFLGEMTGKLDSTMTMELMKPRCG 93
93 VTDNLSYAAAEKISDLFARHRTKMRKRRKFAKQGNKMYKHLSTRLVNWPEHLRSQFG 152
94 VPDVGGFSTF-----PGSPKMKSHITVRIVNY----- 121
153 APCAPSSCGATSQR-----WSSGRPOQAPLITSGSPS---SKGTTTGMAMP LMAQGA 205
122 TPDLPRQVDSAIERALKWBEVTPPLTSRISGEADIMISPAVGEHGDFFPDGFGSL 181
206 RTPLPR---RGAHFDODERMSLSRRGRNLFFVLAHEHGTITGLTHSPAPRALMAPY 262
182 AHAYPRGPFYGDVHFDDKKTLLA-PSGTNLFVLAHEHGTITGLTHSPAPRALMAPY 240
263 KRLGRDA--LISMDVLAVALSYK-PLGGSVAVALPGKLTFTDFETWDSYFQGRAPETQ 319
241 RSTSPANFHLQDDIEGQISLYGAGPSSDAIVVPL-----SVSP---RPER- 285
320 GPKYCHS--SPDAITVDRQOQLYFKGSHWEVAADGNVSEPRPLGERWVGLRPNI EAA 377
286 -PDKCDPALSPDSVSTLRSEVLP-FDRIYFMRSHMNPPEPHILSAFPTLPSDIDAA 343
378 VSLNDGDFYFKGRCMRFRGPPVNLQPOLCRAGGLP---NHPPAALFPPLRLILPK 434
344 EAHNTDSVLIFFKGSQFMVNRGNBVQAGYKGIHTLTGFPPTVKIDAAYEKKKKTYFFV 403
435 GARYVLAARGLOVEPYRSL-QDMGGIP EEVSAGLPRPDGSIIFPRDRIYRLDQAKL 493
404 GDKYRFDETRHVMKGFPRQITDDPGIEPVDAVL-HEFGFFYFFRSGSSQFEED---- 458
494 QATTSGRMATEL-----PMWGC 510
459 ---PNARTVTHILKSNWLLC 476

```

Search completed: June 13, 2005, 18:02:32  
 Job time : 43 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using SW model

Run on: June 13, 2005, 17:14:56 ; Search time 163 Seconds  
(without alignments)  
1233.837 Million cell updates/sec

Title: US-10-791-980-6

Perfect score: 2834  
Sequence: 1 MVARVGLLRALQLILMGHL.....WATLELPMWGCHWANGSALF 520

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database : A\_Geneseq\_15Dec04:\*

1: geneseqp1980s:\*\n2: geneseqp1990s:\*\n3: geneseqp2000s:\*\n4: geneseqp2001s:\*\n5: geneseqp2002s:\*\n6: geneseqp2003as:\*\n7: geneseqp2003bs:\*\n8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 2834  | 100.0       | 520    | AAE14394 | Aae14394 Human mat |
| 2          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 3          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 4          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 5          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 6          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 7          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 8          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 9          | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 10         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 11         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 12         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 13         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 14         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 15         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 16         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 17         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 18         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 19         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 20         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 21         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 22         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 23         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 24         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |
| 25         | 2472  | 87.2        | 520    | AAE14394 | Aae14394 Human pro |

|    |      |      |     |          |                    |
|----|------|------|-----|----------|--------------------|
| 26 | 2472 | 87.2 | 520 | ADA79136 | Ada79136 Human PRO |
| 27 | 2472 | 87.2 | 520 | ADA87275 | Ada87275 Novel hum |
| 28 | 2472 | 87.2 | 520 | ADA16477 | Ada16477 Human PRO |
| 29 | 2472 | 87.2 | 520 | ADA15659 | Ada15659 Novel hum |
| 30 | 2472 | 87.2 | 520 | ADA14632 | Ada14632 Human PRO |
| 31 | 2472 | 87.2 | 520 | ADA18593 | Ada18593 Novel hum |
| 32 | 2472 | 87.2 | 520 | ADA93808 | Ada93808 Human PRO |
| 33 | 2472 | 87.2 | 520 | ADA19704 | Ada19704 Novel hum |
| 34 | 2472 | 87.2 | 520 | ADA13016 | Ada13016 Human PRO |
| 35 | 2472 | 87.2 | 520 | ADA43320 | Ada43320 Novel hum |
| 36 | 2472 | 87.2 | 520 | ADA74270 | Ada74270 Human PRO |
| 37 | 2472 | 87.2 | 520 | ADA24503 | Ada24503 Human PRO |
| 38 | 2472 | 87.2 | 520 | ADA82027 | Ada82027 Human PRO |
| 39 | 2472 | 87.2 | 520 | ADA74590 | Ada74590 Human PRO |
| 40 | 2472 | 87.2 | 520 | ADA85068 | Ada85068 Novel hum |
| 41 | 2472 | 87.2 | 520 | ADA84516 | Ada84516 Novel hum |
| 42 | 2472 | 87.2 | 520 | ADA29772 | Ada29772 Human PRO |
| 43 | 2472 | 87.2 | 520 | ADA80300 | Ada80300 Human PRO |
| 44 | 2472 | 87.2 | 520 | ADA75542 | Ada75542 Human PRO |
| 45 | 2472 | 87.2 | 520 | ADA46767 | Ada46767 Human PRO |

## ALIGNMENTS

|          |  |                            |
|----------|--|----------------------------|
| RESULT 1 | AAE14394   | standard; protein; 520 AA. |
| ID       | AAE14394   | standard; protein; 520 AA. |
| XX       | AAE14394;  |                            |
| AC       | AAE14394;  |                            |
| DT       | 16-OCT-2002 (first entry)  |                            |
| XX       | 16-OCT-2002 (first entry)  |                            |
| DE       | Human matrix metalloproteinase MMP10.                                      |                            |
| XX       | Human matrix metalloproteinase MMP10.                                      |                            |
| XX       | Matrix metalloproteinase; MMP; extracellular matrix; ECM;                  |                            |
| XX       | embryonic development; morphogenesis; reproduction; tissue repair;         |                            |
| XX       | mental disorder; Alzheimer's disease; multiple sclerosis; obesity;         |                            |
| XX       | Parkinson's disease; motor neuron disease; metabolic disease; retinopathy; |                            |
| XX       | type 2 diabetes; cardiovascular; dyslipidaemia; adipogenesis; neuropathy;  |                            |
| XX       | neuropathy; proliferative disease; cancer; psoriasis;                      |                            |
| XX       | prostate hyperplasia; hormonal disorder; alopecia; Crohn's disease;        |                            |
| XX       | central nervous system disorder; CNS; inflammatory condition; arthritis;   |                            |
| XX       | periodontal disease; wound healing; human; MMP10.                          |                            |
| OS       | Homo sapiens.  |                            |
| XX       | Homo sapiens.  |                            |
| XX       | WO200190326-A2.  |                            |
| XX       | 29-NOV-2001.   |                            |
| XX       | 22-MAY-2001; 2001WO-US016563.  |                            |
| XX       | 22-MAY-2001; 2000US-0206119P.  |                            |
| XX       | (PHAA ) PHARMACIA & UPJOHN CO.   |                            |
| XX       | Holmgren E, Kihlen M, Wood T, Ekblom J;                                    |                            |
| XX       | WPI; 2002-083105/11.   |                            |
| XX       | N-PSDB; AAD23965.  |                            |
| XX       | New matrix metalloproteinases (MMP) genes and polypeptides, useful for     |                            |
| XX       | treating diseases or for screening modulators of MMP to treat such         |                            |
| XX       | diseases, e.g. mental disorders, Parkinson's disease, cancers or           |                            |
| XX       | inflammatory conditions.   |                            |
| XX       | Claim 31; Page 60; 94pp; English.  |                            |
| XX       | The invention relates to genes encoding matrix metalloproteinases (MMP).   |                            |
| XX       | The MMP genes are useful for producing MMP polypeptides and for screening  |                            |
| XX       | modulators of MMP. The MMPs are useful for breaking down extracellular     |                            |
| XX       | matrix (ECM), which is essential for processes including embryonic         |                            |

CC development, morphogenesis, reproduction, or tissue repair and  
 CC remodeling. The MMPs are particularly useful for identifying compounds  
 CC that modulate the activity of genes to treat pathologies, e.g. mental  
 CC disorders, Alzheimer's disease, multiple sclerosis, Parkinson's disease  
 CC or motor neuron disease. The MMP polypeptides and genes, as well as their  
 CC modulators, are useful for treating metabolic diseases and disorders  
 CC (e.g. type 2 diabetes, obesity, cardiovascular, dyslipidaemias,  
 CC adipogenesis, retinopathies, neuropathies or nephropathies),  
 CC proliferative diseases and cancers (e.g. breast, colon or lung cancer,  
 CC tumour growth, tumour invasion, psoriasis or prostate hyperplasia),  
 CC hormonal disorders (e.g. male/female hormonal replacement, polycystic  
 CC ovarian syndrome or alopecia), central nervous system (CNS) disorders,  
 CC inflammatory conditions (e.g. Crohn's disease or arthritis), periodontal  
 CC diseases or wound healing. The present sequence is human matrix  
 CC metalloproteinase MMP10

XX  
 SQ Sequence 520 AA;

Query Match 100.0%; Score 2614; DB 5; Length 520;  
 Best Local Similarity 100.0%; Pred. No. 26-251;  
 Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVARVGLLRALQLLMGHLDAQPAERGGOELRKEAEAFLEKYGYNQVPAPISTRS 60  
 DB 1 MVARVGLLRALQLLMGHLDAQPAERGGOELRKEAEAFLEKYGYNQVPAPISTRS 60  
 QY 61 DAIRAFQWVSQLPVSGVLDRAITRQMTPRCGVTDNTNSYAAMERISDLFAHRTKMRK 120  
 DB 61 DAIRAFQWVSQLPVSGVLDRAITRQMTPRCGVTDNTNSYAAMERISDLFAHRTKMRK 120  
 QY 121 KEPAQGNKMYKQHSYRLVNMPEHRSRQFAPCAPSSCGATSGRMSGRPOAPILT 180  
 DB 121 KEPAQGNKMYKQHSYRLVNMPEHRSRQFAPCAPSSCGATSGRMSGRPOAPILT 180  
 QY 121 KEPAQGNKMYKQHSYRLVNMPEHRSRQFAPCAPSSCGATSGRMSGRPOAPILT 180  
 DB 121 KEPAQGNKMYKQHSYRLVNMPEHRSRQFAPCAPSSCGATSGRMSGRPOAPILT 180  
 QY 181 SCSPSSKGTITGMAMPMAQAPWRTPTLPRRGEAFPODERMSLSRRGRNTLVYLAH 240  
 DB 181 SCSPSSKGTITGMAMPMAQAPWRTPTLPRRGEAFPODERMSLSRRGRNTLVYLAH 240  
 QY 241 EIGHTLGLHSPAPALMAYPKRIGRDLMSWDVLAVALSYGKPLGSAVAVQLPGKLF 300  
 DB 241 EIGHTLGLHSPAPALMAYPKRIGRDLMSWDVLAVALSYGKPLGSAVAVQLPGKLF 300  
 QY 301 TDFETWDSYSPQGRRPETQPKYCHSSPDATVDRQOQLYFKGSHFEVAAADVSPR 360  
 DB 301 TDFETWDSYSPQGRRPETQPKYCHSSPDATVDRQOQLYFKGSHFEVAAADVSPR 360  
 QY 301 TDFETWDSYSPQGRRPETQPKYCHSSPDATVDRQOQLYFKGSHFEVAAADVSPR 360  
 DB 301 TDFETWDSYSPQGRRPETQPKYCHSSPDATVDRQOQLYFKGSHFEVAAADVSPR 360  
 QY 361 PLOERWVGLPNIETAAVSLNDGDFYFFKGGRCWFRGPKVWGLPQLCRAGGLPRHPDA 420  
 DB 361 PLOERWVGLPNIETAAVSLNDGDFYFFKGGRCWFRGPKVWGLPQLCRAGGLPRHPDA 420  
 QY 421 ALFPEPLRLILFKGARVYVLARGLQVEPYPPSLQWGGIPREVSALPRPDGSIITF 480  
 DB 421 ALFPEPLRLILFKGARVYVLARGLQVEPYPPSLQWGGIPREVSALPRPDGSIITF 480  
 QY 481 RDDRYWRLDQAKLTATTSGRWATELPMWGCWHAHNSGSLF 520  
 DB 481 RDDRYWRLDQAKLTATTSGRWATELPMWGCWHAHNSGSLF 520

RESULT 2  
 AAB74690  
 ID AAB74690 standard; protein; 520 AA.

XX AAB74690;

DT 12-JUN-2001 (first entry)

DE Human protease and protease inhibitor PPI-23.

KM Human; protease; protease inhibitor; protease and protease inhibitor;  
 KM PPI-23; identification; diagnosis; anti-human immunodeficiency virus; HIV;  
 KM antidiabetic; immunostimulant; immunomodulator; antiinflammatory;  
 KM antitumor; immunosuppressive; nephrotropic; antitumor; thyromimetic;

KM cytostatic; antibacterial; fungicide; protozoacide; antitubercular;  
 KM antitubercular; antiparasitic; virucide; hepatotropic; gene therapy;  
 KM autoimmune disorder; inflammatory disorder; AIDS; Dengue's syndrome;  
 KM severe combined immunodeficiency disease; SCID; Chediak-Higashi syndrome;  
 KM Cushing's disease; Addison's disease; autoimmune thyroiditis; gout;  
 KM Crohn's disease; diabetes mellitus; Good pasture's syndrome; infection;  
 KM Grave's disease; Hashimoto's thyroiditis; Sjogren's syndrome; cancer;  
 KM Werner's syndrome; cell proliferative disorder; arteriosclerosis;  
 KM atherosclerosis; cirrhosis; hepatitis; psoriasis.

XX Homo sapiens.

XX WO200110903-A2.

XX 15-FEB-2001.

XX 09-AUG-2000; 2000WO-US021878.

XX 09-AUG-1999; 99US-0147966P.

XX 21-OCT-1999; 99US-0160807P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Lal P, Tang YT, Bandman O, Baughn MR, Azimzai Y, Lu DAM;  
 PI Yang J,

DR WPI; 2001-2002760/20.

DR N-PSDB; AAF81736.

PT New protease (inhibitors) useful for diagnosis and treatment of  
 PT autoimmune/inflammatory disorders such as acquired immunodeficiency  
 PT syndrome, Cushing's disease, Addison's disease and cell proliferative  
 PT disorders such as cancer.

XX Claim 1; Page 111-112; 134pp; English.

CC AAF8174 to AAF81740 encode the human proteases and protease inhibitors  
 CC (PPI-23) given in AAB74690 to AAB74694. The PPI-23s have activities such  
 CC as: anti-human immunodeficiency virus (HIV); antidiabetic; antitumor;  
 CC immunostimulant; immunomodulator; antiinflammatory; immunosuppressive;  
 CC nephrotropic; antitumor; thyromimetic; cytostatic; antibacterial;  
 CC fungicide; protozoacide; antitubercular; antitubercular;  
 CC virucide; antiparasitic; and hepatotropic. PPI-23 polynucleotide and  
 CC protein sequences can be used in the diagnosis, treatment and prevention  
 CC of autoimmune/inflammatory disorders such as AIDS, Dengue's syndrome,  
 CC severe combined immunodeficiency disease (SCID), Chediak-Higashi  
 CC syndrome, Cushing's disease, Addison's disease, autoimmune thyroiditis,  
 CC Crohn's disease, diabetes mellitus, Good pasture's syndrome, gout,  
 CC Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, Werner's  
 CC syndrome, viral, bacterial, fungal, parasitic, protozoal, and helminthic  
 CC infections and cell proliferative disorder such as arteriosclerosis,  
 CC atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PPI-23  
 CC polynucleotide sequences can be used in somatic or germ-line gene therapy  
 CC and in diagnosis of diseases. They can also be used in generating  
 CC hybridisation probes useful in mapping the naturally occurring genomic  
 CC sequences and in molecular biology techniques

XX Sequence 520 AA;

Query Match 87.2%; Score 2472; DB 4; Length 520;  
 Best Local Similarity 88.9%; Pred. No. 46-218;  
 Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQLLMGHLDAQPAERGGOELRKEAEAFLEKYGYNQVPAPISTRS 60  
 DB 1 MVARVGLLRALQLLMGHLDAQPAERGGOELRKEAEAFLEKYGYNQVPAPISTRS 60  
 QY 61 DAIRAFQWVSQLPVSGVLDRAITRQMTPRCGVTDNTNSYAAMERISDLFAHRTKMRK 120  
 DB 61 DAIRAFQWVSQLPVSGVLDRAITRQMTPRCGVTDNTNSYAAMERISDLFAHRTKMRK 120  
 QY 121 KEPAQGNKMYKQHSYRLVNMPEHRSRQFAPCAPSSCGATSGRMSGRPOAPILT 180  
 DB 121 KEPAQGNKMYKQHSYRLVNMPEHRSRQFAPCAPSSCGATSGRMSGRPOAPILT 180

RESULT 3  
AAB85427  
ID AAB85427 standard; protein; 520 AA

| XX                    | Sequence        | 520 AA:   | 87.2%;                     | Score 2472;       | DB 4; | Length 520;       |
|-----------------------|-----------------|---|----------------------------|-------------------|-------|-------------------|
| Query                 | Match           |   | 88.9%;                     | Pred. No. 4e-218; |       |                   |
| Best Local Similarity | 88.9%;          |   |                            |                   |       |                   |
| Matches               | 472;            | Conservative  | 4;                         | Mismatches        | 33;   | Indels 22; Gaps 5 |
| QY                    | 1               | MVARVLLRLAQLLLMGLHDAQPAERGGQELRKBAEFLKCYVYLNEQVPAKPTSRFS    | 60                         |                   |       |                   |
| DB                    | 1               | MVARVELLRALQLLLMGLHDAQPAERGGQELRKBAEFLKCYVYLNEQVPAKPTSRFS   | 60                         |                   |       |                   |
| QY                    | 61              | DAIRAQWVSQSLPVSGVDLRATLRQMTPRGCVTDITNSYAAMARISDLFAHRTKMRK   | 120                        |                   |       |                   |
| DB                    | 61              | DAIRAQWVSQSLPVSGVDLRATLRQMTPRGCVTDITNSYAAMARISDLFAHRTKMRK   | 120                        |                   |       |                   |
| QY                    | 121             | KREPAKGNKWKYQHLSTYRLVNNPEHLRSRQFAPCAPSSCGATS---QRWS--SGRPP  | 175                        |                   |       |                   |
| DB                    | 121             | KREPAKGNKWKYQHLSTYRLVNNPEHLRSRQFAPCAPSSCGATS---QRWS--SGRPP  | 175                        |                   |       |                   |
| QY                    | 176             | QAPLTSGSSSS-----KQTMTGAMAMLMQAGAPWRFPFLPRRGEAHFODDERWSLSRR  | 225                        |                   |       |                   |
| DB                    | 172             | EAPAT--GPADIRLTLPFGDNDHDGNAFDGPGALAHAFLLPRRGEAHFODDERWSLSRR | 225                        |                   |       |                   |
| QY                    | 230             | RGRNLFPVLAHIEIGHTLGLTTHSPAPRALMAPYRKLGADLLSMDVLAWSLYGKPLGG  | 285                        |                   |       |                   |
| DB                    | 230             | RGRNLFPVLAHIEIGHTLGLTTHSPAPRALMAPYRKLGADLLSMDVLAWSLYGKPLGG  | 285                        |                   |       |                   |
| QY                    | 290             | SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDAITVDROQLYIFKSHFWE    | 345                        |                   |       |                   |
| DB                    | 290             | SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDAITVDROQLYIFKSHFWE    | 345                        |                   |       |                   |
| QY                    | 350             | VAADGNVSEPRPLQERWVGLPNNIDAAVSLNDGPFYFKGGRGCRWRGPKPVWGLPOLC  | 405                        |                   |       |                   |
| DB                    | 350             | VAADGNVSEPRPLQERWVGLPNNIDAAVSLNDGPFYFKGGRGCRWRGPKPVWGLPOLC  | 405                        |                   |       |                   |
| QY                    | 410             | RAGGLFRHPDALFPPLRLRLTFKGAHYVYLARGQLQVEEYPRSLQDMGGLPEEYSGA   | 465                        |                   |       |                   |
| DB                    | 410             | RAGGLFRHPDALFPPLRLRLTFKGAHYVYLARGQLQVEEYPRSLQDMGGLPEEYSGA   | 465                        |                   |       |                   |
| QY                    | 470             | LPEPDSIIFFPRDDRMYRLDQAKQATTSGRWATELPMWGCMHANSGLALF          | 520                        |                   |       |                   |
| DB                    | 470             | LPEPDSIIFFPRDDRMYRLDQAKQATTSGRWATELPMWGCMHANSGLALF          | 520                        |                   |       |                   |
| RESULT 4              |                 |   |                            |                   |       |                   |
| AAU12243              | ID              | AAU12243  | standard; protein; 520 AA. |                   |       |                   |
| XX                    | AAU12243;       |   |                            |                   |       |                   |
| AC                    | 24-OCT-2001     | (first entry)   |                            |                   |       |                   |
| XX                    | Human PR04339   | polypeptide sequence.                                       |                            |                   |       |                   |
| XX                    | Human           |   |                            |                   |       |                   |
| XX                    | Homo sapiens.   |   |                            |                   |       |                   |
| XX                    | MO200140466-A2. |   |                            |                   |       |                   |
| XX                    | 07-JUN-2001.    |   |                            |                   |       |                   |
| XX                    | 01-DEC-2000;    | 2000MO-US032678.  |                            |                   |       |                   |
| XX                    | 01-DEC-1999;    | 99MO-US028301.  |                            |                   |       |                   |
| XX                    | 01-DEC-1999;    | 99MO-US028634.  |                            |                   |       |                   |
| XX                    | 02-DEC-1999;    | 99MO-US028551.  |                            |                   |       |                   |
| XX                    | 02-DEC-1999;    | 99MO-US028564.  |                            |                   |       |                   |
| XX                    | 02-DEC-1999;    | 99MO-US028565.  |                            |                   |       |                   |
| XX                    | 09-DEC-1999;    | 99US-0170262P.  |                            |                   |       |                   |

PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030999.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 05-JAN-2000; 2000WO-US000217.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 03-MAR-2000; 2000US-0187202P.  
 PR 15-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 05-JUN-2000; 2000US-0209832P.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Flvaroff B, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WL, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 DR N-PSDB; AAS21315.  
 XX  
 PT Isolated , secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing PRO  
 PT polypeptides, and detect the presence of mammalian tumors e.g. lung,  
 PT breast, prostate, cervical.  
 XX  
 PS Claim 12; Fig 144; 813p; English.  
 XX  
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane PRO  
 CC polypeptides. The PRO polypeptides are useful to detect other PRO  
 CC polypeptides, to link bioactive molecules to cells expressing PRO  
 CC polypeptides, to modulate biological activities of cells expressing PRO  
 CC polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample. Some  
 CC of the 275 sequences are also useful to stimulate the release of tumour  
 CC necrosis factor-alpha (TNF-alpha) from human blood, the proliferation of  
 CC cells of differentiation of chondrocytes, the proliferation or gene expression in  
 CC pericyte cells, the release of proteoglycans from cartilage, the  
 CC proliferation of inner ear utricular supporting cells or of T-  
 CC lymphocytes, the release of a cytokine from peripheral blood monocytes  
 CC (PBMCs), or the proliferation of endothelial cells. Some of the PRO  
 CC polypeptides may modulate glucose or free fatty acid uptake by skeletal  
 CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor  
 CC VIIA. The PRO polypeptides can be used in assays to identify molecules  
 CC involved in binding interactions. The polynucleotides encoding PRO  
 CC polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy  
 CC  
 XX Sequence 520 AA;  
 SQ

Query Match 87.2%; Score 2472; DB 4; Length 520;  
 Best Local Similarity 88.9%; Pred. No. 4e-218;  
 Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MYARVGLLRALQLLMGHLDAQPAERGGQELRKAEAFLEKYGYLNVQVPAKPTSTPS 60  
 DB 1 MYARVGLLRALQLLMGHLDAQPAERGGQELRKAEAFLEKYGYLNVQVPAKPTSTPS 60  
 QY 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTPRPGCVTDYTNVYAAWAERISDLPARHTKMRK 120  
 DB 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTPRPGCVTDYTNVYAAWAERISDLPARHTKMRK 120  
 QY 121 KRFAQGNKKWKQKHSYTLVNMVPEHLNRQGAFCAPSSCGATS---GRWS--SGRPQP 175  
 DB 121 KRFAQGNKKWKQKHSYTLVNMVPEHLNRQGAFCAPSSCGATS---GRWS--SGRPQP 175  
 QY 176 QAPLTSGSPSS-----KGTTMGWAMPDLMQAQAPWMTPTPLPRGGEAHFDDERWSLSRR 229  
 DB 172 EAPAT--GPADIRLRLFFQGDINDGLGNAFPGGALAHAFPRGGEAHFDDERWSLSRR 229  
 QY 230 RGRNLPVTLAHEIGHTTGLTSPAPRALMAPYKRLGRDALLSWDPVLAQSLYKPLGG 289  
 DB 230 RGRNLPVTLAHEIGHTTGLTSPAPRALMAPYKRLGRDALLSWDPVLAQSLYKPLGG 289  
 QY 290 SVAVQLPQKLPFTDEPTWDSYSPQGRRETPQPKYCHSSFPALITYDROOQLYIFKGSFWE 349  
 DB 290 SVAVQLPQKLPFTDEPTWDSYSPQGRRETPQPKYCHSSFPALITYDROOQLYIFKGSFWE 349  
 QY 350 VAADGNVSEPRPLQERWVGLPNNIEAAVSLNDGDFYFFKGRGCMFRPGKPVWGLPOLC 409  
 DB 350 VAADGNVSEPRPLQERWVGLPNNIEAAVSLNDGDFYFFKGRGCMFRPGKPVWGLPOLC 409  
 QY 410 RAGGLPRHPDPAALFPPPLRLILFKGARVYVLAARGILOVEPYPRSLQDWGIIPEEVSQA 469  
 DB 410 RAGGLPRHPDPAALFPPPLRLILFKGARVYVLAARGILOVEPYPRSLQDWGIIPEEVSQA 469  
 QY 470 LPRPDGSIIFPRDRYRWLDAQILOATTSGRWATELPMWGCWMAHNSGSLF 520  
 DB 470 LPRPDGSIIFPRDRYRWLDAQILOATTSGRWATELPMWGCWMAHNSGSLF 520

## RESULT 5

AAU79810  
 ID AAU79810 standard; protein; 520 AA.  
 AC  
 XX AAU79810;

DT 02-JUL-2002 (first entry)  
 XX

DE Novel matrix metalloproteinase 46798, long form.  
 XX

KW Matrix metalloproteinase; MMP; 46798 long form; heart failure; tumour;  
 KW extracellular matrix degradation; cardiovascular disease; metastasis;  
 KW atherosclerosis; arthritis; nephritis; neurological disease; ischaemia;  
 KW periodontal disease; skin ulceration; liver fibrosis; emphysema; trachea;  
 KW fibrotic lung disease; bacterial infection; viral infection; psoriasis;  
 KW wound healing; chronic injury; autoimmune disorder; angiogenesis;  
 KW tissue invasion.  
 XX

OS Homo sapiens.  
 XX

PN WO200220739-A2.  
 XX

PD 14-MAR-2002.  
 XX

PF 10-SEP-2001; 2001WO-US028260.  
 XX

PR 08-SEP-2000; 2000US-0231136P.  
 XX

PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Curtis RAJ;  
 XX

DR WPI; 2002-351776/38.  
DR N-PSDB; ABK48980.

PT New human matrix metalloproteinase and polymnucleotides useful for  
PT diagnosing and treating atherosclerosis, bacterial and viral infections,  
PT wound healing, chronic injury, traumatic, ischemia and psoriasis.

XX PS Claim 9; Fig 1A-D; 133pp; English.

XX The invention describes a novel isolated human matrix metalloproteinase  
CC (MMP) polypeptide (I) 46798, (I) and the polynucleotide encoding it (II)  
CC are useful for evaluating the efficacy of a treatment of a disorder. (I)  
CC is useful as novel diagnostic target and therapeutic agent for  
CC prognosticating, diagnosing, preventing, inhibiting, alleviating or  
CC curing MMP-related disorders, where (I) functions in normal tissues to  
CC facilitate growth, repair, replacement or renewal of endothelial,  
CC epithelial, and neuronal tissues, by remodeling or degrading  
CC extracellular matrix through or into which new endothelial, epithelial,  
CC or neuronal cells must move, grow or proliferate. Examples of the  
CC disorders involving degradation of the extracellular matrix include  
CC cardiovascular diseases e.g. heart failure and atherosclerosis,  
CC arthritis, nephritis, neurological disease, periodontal disease, skin  
CC ulceration, liver fibrosis, emphysema, fibrotic lung disease, bacterial  
CC and viral infections, wound healing, chronic injury, acute disorder  
CC (autoimmune disorder), traumatic, ischemia, psoriasis, aberrant  
CC angiogenesis and tissue invasion and metastasis by tumour cells. This is  
CC the amino acid sequence of the long form of the novel human matrix  
CC metalloproteinase 46798 described in the invention

XX Sequence 520 AA;

Query Match 87.2%; Score 2472; DB 5; Length 520;

Best Local Similarity 88.9%; Pred. No. 4e-218;

Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

```

QY 1 MYARVGLLRALQLLMGLHDAOPARSGQELRKEAEFLKRYGYNQVPAPISTRPS 60
DB 1 MYARVGLLRALQLLMGLHDAOPARSGQELRKEAEFLKRYGYNQVPAPISTRPS 60
QY 61 DAIRAFQWVSQPVSGVLDRAITLRQMTBRCGVTQDTSNAAVAERISDI.FAHRRTMRK 120
DB 61 DAIRAFQWVSQPVSGVLDRAITLRQMTBRCGVTQDTSNAAVAERISDI.FAHRRTMRK 120
QY 121 KRFPAQGNKMYKQHLSTYRLNMPENLRSRQFAPCAPSSCGATS---GRMS--SGRPOP 175
DB 121 KRFPAQGNKMYKQHLSTYRLNMPENLRSRQFAPCAPSSCGATS---GRMS--SGRPOP 175
QY 176 QAPLTSGSPSS-----KGTTWGMAMPMAQAGAPWRTPLPRGGAHFQDERMSLSRR 229
DB 176 QAPLTSGSPSS-----KGTTWGMAMPMAQAGAPWRTPLPRGGAHFQDERMSLSRR 229
QY 230 RGRNLFVYLAHEIGHTLGTTHSPAPALMAPIYKRLGRALLSMDVLAVQSLYKCPPLG 289
DB 230 RGRNLFVYLAHEIGHTLGTTHSPAPALMAPIYKRLGRALLSMDVLAVQSLYKCPPLG 289
QY 290 SVAVOLPGKLFPTDFTWDSYSPQGRRPETQGPYCHSSPDAITVDROQLYTFKGSHPWE 349
DB 290 SVAVOLPGKLFPTDFTWDSYSPQGRRPETQGPYCHSSPDAITVDROQLYTFKGSHPWE 349
QY 350 VAADGVNSBRPLQERWGLPNIIEAAVSLNDGDFYFFKGRGCRWFGRPKFVWGLPOLC 409
DB 350 VAADGVNSBRPLQERWGLPNIIEAAVSLNDGDFYFFKGRGCRWFGRPKFVWGLPOLC 409
QY 410 RAGGILPRHEDALFPPRLRRLIFKARYYVLARGLQVPEPYPSRLQWGGIPREVSGA 469
DB 410 RAGGILPRHEDALFPPRLRRLIFKARYYVLARGLQVPEPYPSRLQWGGIPREVSGA 469
QY 470 LRPDGSIIFFRDDRWRLDQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520
DB 470 LRPDGSIIFFRDDRWRLDQAKLQATTSGRWATELPMWGCWHAHNSGSLF 520

```

RESULT 6

ABG31460  
ID ABG31460 standard; protein; 520 AA.

XX AC ABG31460;

XX DT 29-NOV-2002 (first entry)

XX DE Human matrix metalloproteinase (MMP) 46798 #2.

XX KW Human; matrix metalloproteinase; MMP 46798; cell proliferation disorder;  
KW cell differentiation disorder; carcinoma; sarcoma; leukaemia;  
KW breast cancer; lung cancer; neurological disorder; schizophrenia;  
KW ischaemia; infarction; Parkinson's disease; Huntington's disease;  
KW inflammatory disorder; Crohn's disease; immune disorder; arthritis;  
KW diabetes mellitus; cardiovascular disorder; reestenosis; tachycardia;  
KW rheumatic heart disease; motility disorder; developmental disorder;  
KW lung disorder; chronic bronchitis; pulmonary congestion; oedema;  
KW blood disorder; blood clotting disorder; cytosatic; immunomodulator;  
KW anti-inflammatory; cardiac; antiparkinsonian; nootropic; thrombolytic;  
KW neuroprotective; antidiabetic; antirheumatic; antiarthritic; vasotropic;  
KW enzyme.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Domain 39..225

XX FT Domain /label= Peptidase\_M10\_domain

XX FT Domain 328..371

XX FT Domain /label= Haemopexin-like\_domain

XX FT Domain 373..416

XX FT Domain /label= Haemopexin-like\_domain

XX FT Domain 418..464

XX FT Domain /label= Haemopexin-like\_domain

XX FT Domain 466..510

XX FT Domain /label= Haemopexin-like\_domain

XX PN WO20026670-A2.

XX PD 29-AUG-2002.

XX PF 16-JAN-2002; 2002WO-US001546.

XX PR 16-JAN-2001; 2001US-0262252P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Curtis RAJ, Lora JM;

XX PT WPI; 2002-674955/72.

XX DR N-PSDB; ABK91096.

XX PS Claim 17; Page 113-114; 117pp; English.

PT New human matrix metalloproteinase nucleic acid and polypeptide  
PT molecules, designated 46798, useful for diagnosing, preventing or  
PT treating cancers, ischemia, Parkinson's disease heart disease or edema.  
XX

XX The present invention relates to the isolation of novel human matrix  
CC metalloproteinases (MMP), designated 46798, and the polynucleotide  
CC sequences encoding them. The MMP 46798 polypeptide and polynucleotide  
CC sequences are useful for diagnosing, preventing, alleviating or treating  
CC metalloproteinase-associated disorders such as cell proliferation and/or  
CC differentiation disorders (e.g. carcinoma, sarcoma, leukaemia, breast  
CC cancer, or lung cancer), neurological disorders (e.g. schizophrenia,  
CC ischaemia, infarction, Parkinson's disease or Huntington's disease),  
CC inflammatory disorders (e.g. Crohn's disease) and immune disorders (e.g.  
CC diabetes mellitus or arthritis), cardiovascular disorders (e.g.  
CC reestenosis, tachycardia or rheumatic heart disease), motility disorders,  
CC developmental disorders, lung disorders (e.g. chronic bronchitis,  
CC pulmonary congestion or oedema), and blood/blood clotting disorders. The  
CC present sequence represents human MMP 46798 #2

XX Sequence 520 AA;

Query Match 87.2%; Score 2472; DB 5; Length 520;  
 Best Local Similarity 88.9%; Pred. No. 4e-218;  
 Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVAAGLLRLAQLLLMGLHDAOPAEKGGGELREKAEAFLEKGYLYNEQVPAKPTSTRFS 60  
 DB 1 MVAAGLLRLAQLLLMGLHDAOPAEKGGGELREKAEAFLEKGYLYNEQVPAKPTSTRFS 60  
 QY 61 DAIAFQWVSQLPVSGVLDRAITLQMTTRPCGVDTINSYAAMERISDLFARHTKMRK 120  
 DB 61 DAIAFQWVSQLPVSGVLDRAITLQMTTRPCGVDTINSYAAMERISDLFARHTKMRK 120  
 QY 121 KRFAKQGNKYKQHLSTRLVNMPEHLRSROFGAPCAPSSCGATS--QKWS--SGRPOP 175  
 DB 121 KRFAKQGNKYKQHLSTRLVNMPEHLRSROFGAPCAPSSCGATS--QKWS--SGRPOP 175  
 QY 176 QAPLTSSGSS-----KGTITMGWAMPILMAOGAPWPTPLPRGEAHFPODERWLSRR 229  
 DB 172 EAPAT--GPADIRLTFPGQDHNDGLGNAFGCGALAHAFPRGEAHFPODERWLSRR 229  
 QY 230 RGRNLFVYLAEHIGHTGLTHSPAPALMAPYKRLGRDALLSMDVLAQSLYKPKLGG 289  
 DB 230 RGRNLFVYLAEHIGHTGLTHSPAPALMAPYKRLGRDALLSMDVLAQSLYKPKLGG 289  
 QY 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDATVDRQOOLYFKGSHWE 349  
 DB 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDATVDRQOOLYFKGSHWE 349  
 QY 350 VAADVNSEPRPQOERWVGIPNIEAAVSLNDGDFPFKGRGCRFRGPKVWGIPOLC 409  
 DB 350 VAADVNSEPRPQOERWVGIPNIEAAVSLNDGDFPFKGRGCRFRGPKVWGIPOLC 409  
 QY 410 RAGGLPRHPDALFFPPLRLILFKGARVYLARGLQVEPYPPSLQDWGIPPEVSGA 469  
 DB 410 RAGGLPRHPDALFFPPLRLILFKGARVYLARGLQVEPYPPSLQDWGIPPEVSGA 469  
 QY 470 LRRPDGSIIFRRDRYWRLDQAKLQATTSGRWATELPMWGCNANSGSALF 520  
 DB 470 LRRPDGSIIFRRDRYWRLDQAKLQATTSGRWATELPMWGCNANSGSALF 520

RESULT 7  
 ID ABO17687 standard; protein; 520 AA.  
 AC ABO17687;  
 XX 26-AUG-2003 (first entry)  
 DT XX  
 DE Novel human secreted and transmembrane protein PRO4339.  
 XX  
 KW Human; secreted and transmembrane protein; PRO; anti-inflammatory;  
 KW antiarteriosclerotic; cardiant; anti-fertility; anti-HIV; cytostatic;  
 KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;  
 KW TNF-alpha release; cell proliferation; cell differentiation;  
 KW gene expression modulator; proteoglycan release; cytokine release;  
 KW tumour; inflammatory disease; organ failure; atherosclerosis;  
 KW cardiac injury; infertility; birth defect; premature aging; AIDS;  
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;  
 KW chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;  
 KW bioreactor; tissue typing.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003032156-A1.  
 XX  
 PD 13-FEB-2003.  
 XX  
 PF 06-MAY-2002; 2002US-00140474.  
 XX  
 PR 31-MAR-1997; 97WO-US005230.  
 PR 12-JUN-1998; 98WO-US012456.

PR 14-JUL-1998; 98WO-US014552.  
 PR 26-AUG-1998; 98WO-US017888.  
 PR 10-SEP-1998; 98WO-US018824.  
 PR 14-SEP-1998; 98WO-US019093.  
 PR 14-SEP-1998; 98WO-US019094.  
 PR 14-SEP-1998; 98WO-US019177.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98WO-US019437.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 29-OCT-1998; 98WO-US022991.  
 PR 29-OCT-1998; 98WO-US022992.  
 PR 20-NOV-1998; 98WO-US024855.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 05-JAN-1999; 99WO-US000106.  
 PR 08-MAR-1999; 99WO-US0005028.  
 PR 10-MAR-1999; 99WO-US0005190.  
 PR 20-APR-1999; 99WO-US0008615.  
 PR 14-MAY-1999; 99WO-US010733.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 08-SEP-1999; 99WO-US020594.  
 PR 13-SEP-1999; 99WO-US020944.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 15-SEP-1999; 99WO-US021547.  
 PR 05-OCT-1999; 99WO-US023089.  
 PR 29-NOV-1999; 99WO-US028214.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 01-DEC-1999; 99WO-US028409.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 02-DEC-1999; 99WO-US028634.  
 PR 02-DEC-1999; 99WO-US028651.  
 PR 02-DEC-1999; 99WO-US028654.  
 PR 02-DEC-1999; 99WO-US028655.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030999.  
 PR 22-DEC-1999; 99WO-US030720.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005064.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005746.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 10-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007337.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020711.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 01-DEC-2000; 2000WO-US030873.  
 PR 10-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001US-00796498.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-MAR-2001; 2001WO-US006566.  
 PR 09-MAR-2001; 2001US-00802706.



PR 14-MAR-2001; 2001US-00808689.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 05-APR-2001; 2001US-00828366.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 10-MAY-2001; 2001US-00854280.  
 PR 18-MAY-2001; 2001US-00860216.  
 PR 25-MAY-2001; 2001US-0086034.  
 PR 25-MAY-2001; 2001US-00866034.  
 PR 01-JUN-2001; 2001US-0087035.  
 PR 01-JUN-2001; 2001US-0087035.  
 PR 05-JUN-2001; 2001US-00874503.  
 PR 14-JUN-2001; 2001US-00882636.  
 PR 19-JUN-2001; 2001US-00886342.  
 PR 20-JUN-2001; 2001US-0089692.  
 PR 21-JUN-2001; 2001US-00887879.  
 PR 22-JUN-2001; 2001US-008920116.  
 PR 29-JUN-2001; 2001US-008920116.  
 PR 09-JUL-2001; 2001US-008920116.  
 PR 18-JUL-2001; 2001US-00908827.  
 PR 06-AUG-2001; 2001US-00924419.  
 PR 09-AUG-2001; 2001US-00927796.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 19-DEC-2001; 2001US-00028072.  
 XX (GENTH ) GENENTECH INC.  
 XX Baker KP, Betesini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PO, Gurney AL, Sherwood Z;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang S;  
 DR WPI; 2003-341980/32.  
 DR N-PSDB; ACD33924.  
 XX  
 PT New secreted and transmembrane PRO nucleic acids, for treating  
 PT inflammation, organ failure, atherosclerosis, cardiac injury,  
 PT infertility, birth defects, premature aging, acquired immunodeficiency  
 PT syndrome (AIDS), or cancer.  
 XX  
 PS Claim 12; Fig 144; 660pp; English.  
 CC The invention describes an isolated nucleic acid (1) comprising, or which  
 CC has 80 % sequence identity to, or the full-length coding sequence of, one  
 CC of 275 nucleotide sequences, and which encodes a corresponding  
 CC polypeptide selected from 275 amino acid sequences, where all sequences  
 CC are given in the specification. The polypeptide encoded by (1) is used to  
 CC detect PRO polypeptides, link a bioactive molecule to a cell expressing a  
 CC PRO polypeptide, modulate a biological activity of a cell, stimulate the  
 CC release of tumour necrosis factor (TNF)-alpha from human blood, modulate  
 CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit  
 CC the proliferation or differentiation of cells or gene expression.  
 CC stimulate the release of proteoglycans, stimulate the release of cytokine  
 CC from peripheral blood mononuclear cells, inhibit the binding of A-peptide  
 CC to factor VIIa, or detect the presence of tumour in a mammal. The nucleic  
 CC acid and polypeptide encoded by it, are useful for treating inflammatory  
 CC diseases, organ failure, atherosclerosis, cardiac injury, infertility,  
 CC birth defects, premature aging, acquired immunodeficiency syndrome  
 CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as  
 CC hybridisation probes, in chromosome and gene mapping, and in generating  
 CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,  
 CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.  
 CC This is the amino acid sequence of a novel human secreted and  
 CC transmembrane PRO polypeptide  
 XX  
 SO Sequence 520 AA;  
 Query Match 87.2%; Score 2472; DB 6; Length 520;  
 Best Local Similarity 88.9%; Pred. No. 4e-218;  
 Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 61 DAIRAFOWWSQPVSGVLDRAFLRQMTREPCCVTDNYSAAWAERISDLFAHRTMKRK 120  
 DB 61 DAIRAFOWWSQPVSGVLDRAFLRQMTREPCCVTDNYSAAWAERISDLFAHRTMKRK 120  
 QY 121 KRPFAQGNKMYKQHLSTYLVNMPHLSRQFQCAPPSSCATS--ORWG--SGRPOP 175  
 DB 121 KRPFAQGNKMYKQHLSTYLVNMPHLSRQFQCAPPSSCATS--ORWG--SGRPOP 175  
 QY 176 QAPLTSGSPSS-----KGTMTGWAMPPLMAQCAPWRPTEFLRRRGPAHFDQDERMSLSR 229  
 DB 172 EAPAT--GPADIRLTFPGDHNHNGJGNMFDGFGALAHAFPRRGSAHFDQDERMSLSR 229  
 QY 220 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAFYKRLGSDALUSMDVLAQSLYKPLCG 289  
 DB 220 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAFYKRLGSDALUSMDVLAQSLYKPLCG 289  
 QY 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAFYKRLGSDALUSMDVLAQSLYKPLCG 289  
 DB 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAFYKRLGSDALUSMDVLAQSLYKPLCG 289  
 QY 290 SVAVOLPGKLPDPEFTWDSYSFQGRPETQGRKCHSSFDATVDRQOOLYFKGSHFWE 349  
 DB 290 SVAVOLPGKLPDPEFTWDSYSFQGRPETQGRKCHSSFDATVDRQOOLYFKGSHFWE 349  
 QY 290 SVAVOLPGKLPDPEFTWDSYSFQGRPETQGRKCHSSFDATVDRQOOLYFKGSHFWE 349  
 DB 290 SVAVOLPGKLPDPEFTWDSYSFQGRPETQGRKCHSSFDATVDRQOOLYFKGSHFWE 349  
 QY 350 VAADNVSEPRPLQGRWGLPNTIRAAVSLNDGDFYFKGGRCMRFRGKPVWGLPOLC 409  
 DB 350 VAADNVSEPRPLQGRWGLPNTIRAAVSLNDGDFYFKGGRCMRFRGKPVWGLPOLC 409  
 QY 410 RAGGLPRHPDALFPPLRLILFKGARYYVLARGGLQVEPYPSLQDMGIPPEVSGA 469  
 DB 410 RAGGLPRHPDALFPPLRLILFKGARYYVLARGGLQVEPYPSLQDMGIPPEVSGA 469  
 QY 470 LRPDGSIIFFPDRIYWRIDQAKLQATTSGRWATELPMWGMHANSGLALF 520  
 DB 470 LRPDGSIIFFPDRIYWRIDQAKLQATTSGRWATELPMWGMHANSGLALF 520  
 RESULT 8  
 ABU80941  
 ID ABU80941 standard; protein; 520 AA.  
 XX  
 AC ABU80941;  
 DT 23-JUN-2003 (first entry)  
 XX  
 DE Human PRO polypeptide #72.  
 XX  
 KW Human; PRO polypeptide; secreted and transmembrane protein;  
 KW anti-PRO antibody; diagnostic assay; gene expression; diabetes;  
 KW bone disorder; cartilage disorder; rheumatoid arthritis; obesity;  
 KW sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;  
 KW hearing loss; coagulation disorder; stroke; heart attack; cardiast;  
 KW antidiabetic; anorectic; vulnary; antirheumatic; osteopathic;  
 KW antirheumatic; auditory; cerebroprotective; angiogenic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US200304311-A1.  
 XX  
 PD 02-JAN-2003.  
 XX  
 PF 19-DEC-2001; 2001US-00028072.  
 XX  
 XX 18-JUN-1997; 97US-0049911P.  
 XX 26-AUG-1997; 97US-0056974P.  
 XX 17-SEP-1997; 97US-0059113P.  
 XX 17-SEP-1997; 97US-0059115P.  
 XX 17-SEP-1997; 97US-0059117P.  
 XX 17-SEP-1997; 97US-0059122P.  
 XX 18-SEP-1997; 97US-0059263P.  
 XX 19-SEP-1997; 97US-0059352P.  
 XX 19-SEP-1997; 97US-0059386P.  
 XX 24-SEP-1997; 97US-0059836P.  
 XX 17-OCT-1997; 97US-0062250P.  
 XX 17-OCT-1997; 97US-0062285P.

[illegible]

Db 290 SVAVOLPGKLTFTDEFTWDSYSPQGRRPETQGPKYCHSSFDALTVDROOQLYFKGSHWE 349  
Qy 350 VAADGNVEBPRLQGRWVGLPNPINEAAVSLNDGDFYFPGKGRCMRFRGPKFVMGLPOLC 409  
Db 350 VAADGNVEBPRLQGRWVGLPNPINEAAVSLNDGDFYFPGKGRCMRFRGPKFVMGLPOLC 409  
Qy 410 RAGGLPRHPDALFPPEPLRRLILFKGARYYVLARGGLQVEPYPRSLQDMGIPREVSQA 469  
Db 410 RAGGLPRHPDALFPPEPLRRLILFKGARYYVLARGGLQVEPYPRSLQDMGIPREVSQA 469  
Qy 470 LPRPGSIIFFRDDRYWRDLQAKQATTSGRWATELPMWGMCHANSGLP 520  
Db 470 LPRPGSIIFFRDDRYWRDLQAKQATTSGRWATELPMWGMCHANSGLP 520

RESULT 9  
ABU66641  
ID ABU66641 standard; protein; 520 AA.  
XX  
AC ABU66641;  
XX  
DT 23-MAY-2003 (first entry)  
XX  
DE Human PRO polypeptide #72.  
XX  
KW Human; PRO polypeptide; secreted and transmembrane protein;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;  
KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN US2003036180-A1.  
XX  
PD 20-FEB-2003.  
XX  
PF 09-MAY-2002; 2002US-00143114.  
XX  
PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 20-NOV-1998; 98WO-US024855.  
PR 05-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US012252.  
PR 02-JUN-1999; 99WO-US010733.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028651.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030939.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 11-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US008439.  
PR 30-MAR-2000; 2000WO-US013705.  
PR 17-MAY-2000; 2000WO-US014042.  
PR 22-MAY-2000; 2000WO-US014941.  
PR 30-MAY-2000; 2000WO-US015264.  
PR 02-JUN-2000; 2000WO-US020710.  
PR 28-JUN-2000; 2000WO-US022031.  
PR 11-AUG-2000; 2000WO-US023522.  
PR 23-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00860216.  
PR 18-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.

XX  
XX (GENTH ) GENENTECH INC.  
XX  
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
XX Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,  
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z,  
XX MPI; 2003-332040/31.  
XX DR N-PSDB; ACN03674.

PT New secreted and transmembrane PRO nucleic acids, useful for gene  
therapy, in chromosome and gene mapping, as chromosome markers, in tissue  
typing, and in chromosome identification.

XX Claim 12; Fig 144; 660pp; English.

XX The present invention relates to the isolation of novel human PRO  
CC polypeptides, and the polynucleotide sequences encoding them. The PRO  
CC polypeptides are secreted and transmembrane proteins. The PRO  
CC polypeptides are useful for detecting other PRO polypeptides, for linking  
CC bioactive molecules to cells expressing PRO polypeptides, for modulating  
CC biological activities of cells expressing PRO polypeptides, and for  
CC identifying agonists or antagonists. The PRO polypeptides are useful for  
CC human blood, for stimulating the release of tumor necrosis factor (TNF)-alpha from  
CC chondrocytes, and detecting the presence of tumors. The polynucleotide  
CC sequences encoding PRO polypeptides are useful as hybridization probes,  
CC in chromosome and gene mapping, in the generation of antisense RNA and  
CC DNA, in the preparation of PRO polypeptides, for generating transgenic  
CC animals or knockout animals, for the genetic analysis of individuals with  
CC genetic disorders, and in gene therapy. AB06570-AB06644 represent the  
CC human PRO polypeptides of the invention. Note: The sequence data for this  
CC patent was obtained in electronic format directly from the USPTO web site  
CC at seqdata.uspto.gov/psipdsidentry.html  
XX

XX Sequence 520 AA;

Query Match 87.2%; Score 2472; DB 6; Length 520;  
Best Local Similarity 88.9%; Pred. No. 4e-218;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQLLMGHLDAPERRGOELRKEAEAELEKYGLNQPAPSTPS 60  
DB 1 MVARVGLLRALQLLMGHLDAPERRGOELRKEAEAELEKYGLNQPAPSTPS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAITLROMTREPCVDTNYSAAWAERISDLFAHRTQMRK 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAITLROMTREPCVDTNYSAAWAERISDLFAHRTQMRK 120  
QY 121 KRFAQGNKWKYKOHLSYLVNMPPEHLRSQFAPCAPSSCCGANS--GRS--SGRDP 175  
DB 121 KRFAQGNKWKYKOHLSYLVNMPPEHLRSQFAPCAPSSCCGANS--GRS--SGRDP 175  
QY 176 QAPLTSGSPSS-----KQTTMGWAMPPLMAQGAFWRTPLPRGEAHFDDERSLSRR 229  
DB 172 EAPAT--GPADIRLTFPGGDHNDGIGNAPDGGALAAFLPRGEAHFDDERSLSRR 229  
QY 230 RGRNLFVVLAEHIGHTLGLTSPAPRALMAPYKRLGRDALLSMDVLAVALGKPLGG 289  
DB 230 RGRNLFVVLAEHIGHTLGLTSPAPRALMAPYKRLGRDALLSMDVLAVALGKPLGG 289  
QY 290 SVAVOLPEKLTFTDEFTWVSYPQGRRPETQPKYCHSFDAITVDROQLYIFGSHFWE 349  
DB 290 SVAVOLPEKLTFTDEFTWVSYPQGRRPETQPKYCHSFDAITVDROQLYIFGSHFWE 349  
QY 350 VAADGNVSEPRPLQERWGLPNNIEAAVSLNDGDFYFPGKRCMRFRGPKVWGLPOLC 409  
DB 350 VAADGNVSEPRPLQERWGLPNNIEAAVSLNDGDFYFPGKRCMRFRGPKVWGLPOLC 409  
QY 410 RAGGLPRPDAALFPPPLRLILFKGARYVVLARGLQVEBPYRSLQDWGGLPEEVSQA 469  
DB 410 RAGGLPRPDAALFPPPLRLILFKGARYVVLARGLQVEBPYRSLQDWGGLPEEVSQA 469  
QY 470 LPRPDGSIIFPRDRRYMLDQAKIQAATTSQWAMELPMWGMWANSALF 520  
DB 470 LPRPDGSIIFPRDRRYMLDQAKIQAATTSQWAMELPMWGMWANSALF 520

RESULT 10  
ABUS9722  
ID ABUS9722 standard; protein; 520 AA.  
XX  
AC ABUS9722;

XX 13-MAY-2003 (first entry)  
DT Novel secreted and transmembrane protein PRO4339.  
XX  
XX Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
KW cardiac insufficiency disorder; cancer; tumour; immune response;  
KW adrenal cortical capillary endothelial growth; c-fos induction;  
KW vascular endothelial growth factor inhibition; VEGF inhibition;  
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;  
KW retinal neurons cell survival; rod photoreceptor cell survival;  
KW retinal disorder; retinitis pigmentosa; kidney disorder;  
KW mammalian kidney mesangial cell proliferation; Berger disease;  
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
KW chondrocyte redifferentiation; sports injury; arthritis.  
XX  
XX Homo sapiens.  
XX  
XX US2003017563-A1.  
XX  
XX 23-JAN-2003.  
XX  
XX 07-MAY-2002; 2002US-00140808.  
XX  
XX 31-MAR-1997; 97WO-US005230.  
XX 12-JUN-1998; 98WO-US012456.  
XX 14-JUL-1998; 98WO-US014552.  
XX 28-AUG-1998; 98WO-US017888.  
XX 10-SEP-1998; 98WO-US018824.  
XX 14-SEP-1998; 98WO-US018824.  
XX 14-SEP-1998; 98WO-US019093.  
XX 14-SEP-1998; 98WO-US019094.  
XX 14-SEP-1998; 98WO-US019177.  
XX 16-SEP-1998; 98WO-US019330.  
XX 17-SEP-1998; 98WO-US019437.  
XX 07-OCT-1998; 98WO-US021141.  
XX 29-OCT-1998; 98WO-US022891.  
XX 29-OCT-1998; 98WO-US022891.  
XX 20-NOV-1998; 98WO-US024855.  
XX 01-DEC-1998; 98WO-US025108.  
XX 05-JAN-1999; 99WO-US000106.  
XX 08-MAR-1999; 99WO-US005028.  
XX 10-MAR-1999; 99WO-US005190.  
XX 20-APR-1999; 99WO-US008615.  
XX 14-MAY-1999; 99WO-US010733.  
XX 02-JUN-1999; 99WO-US012252.  
XX 01-SEP-1999; 99WO-US020111.  
XX 08-SEP-1999; 99WO-US020594.  
XX 13-SEP-1999; 99WO-US020944.  
XX 15-SEP-1999; 99WO-US021090.  
XX 15-SEP-1999; 99WO-US021547.  
XX 05-OCT-1999; 99WO-US023089.  
XX 29-NOV-1999; 99WO-US028214.  
XX 30-NOV-1999; 99WO-US028313.  
XX 30-NOV-1999; 99WO-US028409.  
XX 01-DEC-1999; 99WO-US028301.  
XX 01-DEC-1999; 99WO-US028334.  
XX 02-DEC-1999; 99WO-US028551.  
XX 02-DEC-1999; 99WO-US028564.  
XX 02-DEC-1999; 99WO-US028565.  
XX 16-DEC-1999; 99WO-US030095.  
XX 20-DEC-1999; 99WO-US030911.  
XX 20-DEC-1999; 99WO-US030939.  
XX 22-DEC-1999; 99WO-US030720.  
XX 30-DEC-1999; 99WO-US031243.  
XX 30-DEC-1999; 99WO-US031274.  
XX 05-JAN-2000; 2000WO-US000219.  
XX 06-JAN-2000; 2000WO-US000277.  
XX 06-JAN-2000; 2000WO-US000376.  
XX 11-FEB-2000; 2000WO-US003565.  
XX 18-FEB-2000; 2000WO-US004341.  
XX 18-FEB-2000; 2000WO-US004342.  
XX 22-FEB-2000; 2000WO-US004414.  
XX 24-FEB-2000; 2000WO-US004914.



XX Human secreted/transmembrane protein (PRO) #72.  
 DE  
 XX  
 KW Human; PRO; secreted protein; transmembrane protein; tumour; cytostatic;  
 KW gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;  
 KW proteoglycan; cartilage; cytokine; peripheral blood mononuclear cell;  
 KW PBMG; glucose uptake; FFA; skeletal muscle cell; adipocyte cell;  
 KW chondrocyte cell proliferation; chondrocyte cell differentiation;  
 KW pericyte cell; inner ear utricular supporting cell; T-lymphocyte cell;  
 KW endothelial cell; A-peptide; factor VIIA.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2003036179-A1.  
 PN  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 10-MAY-2002; 2002US-00142431.  
 XX  
 XX 31-MAR-1997; 97WO-US005230.  
 PR 12-JUN-1998; 98WO-US012456.  
 PR 14-JUL-1998; 98WO-US014552.  
 PR 28-AUG-1998; 98WO-US017888.  
 PR 10-SEP-1998; 98WO-US018824.  
 PR 14-SEP-1998; 98WO-US019093.  
 PR 14-SEP-1998; 98WO-US019094.  
 PR 16-SEP-1998; 98WO-US019177.  
 PR 17-SEP-1998; 98WO-US019330.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 29-OCT-1998; 98WO-US022991.  
 PR 29-OCT-1998; 98WO-US022992.  
 PR 20-NOV-1998; 98WO-US024855.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 05-JAN-1999; 99WO-US000106.  
 PR 08-MAR-1999; 99WO-US005028.  
 PR 10-MAR-1999; 99WO-US005190.  
 PR 20-APR-1999; 99WO-US008615.  
 PR 14-MAY-1999; 99WO-US010733.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 08-SEP-1999; 99WO-US020594.  
 PR 13-SEP-1999; 99WO-US020944.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 15-SEP-1999; 99WO-US021547.  
 PR 05-OCT-1999; 99WO-US023069.  
 PR 29-NOV-1999; 99WO-US028214.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 30-NOV-1999; 99WO-US028409.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 02-DEC-1999; 99WO-US028634.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 02-DEC-1999; 99WO-US028564.  
 PR 16-DEC-1999; 99WO-US028565.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030939.  
 PR 22-DEC-1999; 99WO-US030720.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005054.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005746.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 10-MAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000US-00796498.  
 PR 28-FEB-2001; 2001US-00828366.  
 PR 28-FEB-2001; 2001US-00828366.  
 PR 01-MAR-2001; 2001WO-US006560.  
 PR 09-MAR-2001; 2001US-00802706.  
 PR 14-MAR-2001; 2001US-00808689.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 05-APR-2001; 2001US-00828366.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 18-MAY-2001; 2001US-00860216.  
 PR 25-MAY-2001; 2001US-00866028.  
 PR 25-MAY-2001; 2001US-00866034.  
 PR 25-MAY-2001; 2001WO-US017092.  
 PR 01-JUN-2001; 2001US-00872035.  
 PR 05-JUN-2001; 2001WO-US017800.  
 PR 14-JUN-2001; 2001US-00882536.  
 PR 19-JUN-2001; 2001US-00882536.  
 PR 20-JUN-2001; 2001US-00886342.  
 PR 21-JUN-2001; 2001WO-US019692.  
 PR 22-JUN-2001; 2001WO-US020116.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 18-JUL-2001; 2001US-00908827.  
 PR 06-AUG-2001; 2001US-00924419.  
 PR 09-AUG-2001; 2001US-00927966.  
 PR 16-AUG-2001; 2001US-00931836.  
 PR 19-DEC-2001; 2001US-00028072.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
 PI Gertlissen WE, Goddard A, Godowski PJ, Gurney AL, Sherwood S,  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z,  
 XX WPI; 2003-466355/44.  
 DR N-PSDB; ACD41866.  
 DR  
 XX  
 PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO114 or  
 PT PRO4978, useful in molecular biology, chromosome and gene mapping, in  
 PT generating antisense RNA and DNA, and in gene therapy.  
 XX  
 PS Claim 12; Fig 144; 659pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising at least 80%  
 CC sequence identity to a PRO (secreted and transmembrane protein) cDNA  
 CC comprising a nucleic acid (a) encoding a PRO polypeptide, or its  
 CC extracellular domain (with or without its associated signal peptide),  
 CC which comprises any of the 275 120-850 residue amino acid sequences,  
 CC given in the specification; (b) comprising any of the 275 300-3500  
 CC nucleotide sequences, given in the specification; or (c) comprising the  
 CC full-length coding sequence of the nucleotide sequences given in the  
 CC specification, or of the DNA deposited under any of the American Type  
 CC Culture Collection (ATCC) Accession Numbers listed in the specification.  
 CC Also included are a vector comprising the novel nucleic acid, a host cell  
 CC comprising the vector, producing a PRO polypeptide, the isolated PRO

polypeptides detailed above, a chimaeric molecule comprising the PRO polypeptide of fused to a heterologous amino acid sequence, an anti-PRO antibody, detecting a PRO polypeptide in a sample suspected of containing the PRO polypeptide, linking a bioactive molecule to a cell expressing a PRO polypeptide, modulating at least one biological activity of a cell expressing a PRO polypeptide, stimulating the release of tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) from human blood, (or proteoglycans from cartilage or cytokine from peripheral blood mononuclear cells (PBMC)), modulating the uptake of glucose or FFA by skeletal muscle cells or adipocyte cells, stimulating the proliferation or differentiation of chondrocyte cells (or proliferation of or gene expression in pericyte cells), stimulating the proliferation of inner ear utricular supporting cells (or of T-lymphocyte cells, or of endothelial cells), inhibiting the binding of A-peptide to factor VIIa, or differentiation of adipocyte cells, detecting the presence of a tumour in a mammal and an oligonucleotide probe derived from any of the nucleotide sequences given in the specification. The polynucleotide is useful in molecular biology, including uses as hybridisation probes, in chromosome and gene mapping, in generating antisense RNA and DNA, and in gene therapy. The polynucleotide may also be used in preparing PRO polypeptides by recombinant techniques, and in generating either transgenic animals or knock-out animals which, in turn, are useful in the development and screening of therapeutically useful reagents. The PRO polypeptide or the antibody is used in preparing a medicament for treating a condition responsive to the polypeptide or antibody, such as tumours, and in various diagnostic assays. The present sequence represents a PRO polypeptide

Sequence 520 AA:

Query Match 87.2%; Score 2472; DB 6; Length 520;  
Best Local Similarity 88.9%; Pred. No. 4e-218;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MAAVGLLRAQLQIMGHLDQAPARRGGELKKEAELEKGYINEQVPAPISTRS 60  
DB 1 MAAVGLLRAQLQIMGHLDQAPARRGGELKKEAELEKGYINEQVPAPISTRS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDNSTAAAEERISDFAHRTMRK 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAITLRQMTTRPCGVTDNSTAAAEERISDFAHRTMRK 120  
QY 121 KRFAGQNKMYKQHLSTYLVNMPHLSRQFGAPCAPSSCGATS--GRWS--SGRPOP 175  
DB 121 KRFAGQNKMYKQHLSTYLVNMPHLSRQFGAPCAPSSCGATS--GRWS--SGRPOP 175  
QY 176 CAPPLTSGSSSS-----KGTITMGAMMPMAOGAPKPTFPLPRRGAHFDODBSLSRR 229  
DB 176 CAPPLTSGSSSS-----KGTITMGAMMPMAOGAPKPTFPLPRRGAHFDODBSLSRR 229  
QY 230 RGRNLFVYLAHETGHTLSTSPAPRALMAPYKRLGRDALISMVDVLAVOSLYGKPLDG 289  
DB 230 RGRNLFVYLAHETGHTLSTSPAPRALMAPYKRLGRDALISMVDVLAVOSLYGKPLDG 289  
QY 290 SVAVALPGKLTFTDFETWDSYSFQGRRPETQGPXYCHSSFDATTVDRQQLYIFKSGHFE 349  
DB 290 SVAVALPGKLTFTDFETWDSYSFQGRRPETQGPXYCHSSFDATTVDRQQLYIFKSGHFE 349  
QY 350 VAADGNVSEPRPLQERWVGLPNIEAAVSLNDGDFYFPFKGRCWRFRGPKVWGIPQLC 409  
DB 350 VAADGNVSEPRPLQERWVGLPNIEAAVSLNDGDFYFPFKGRCWRFRGPKVWGIPQLC 409  
QY 410 RAGGLPRHPDALFPPLRLTLFGARYYVLARGLQVEPYPSLQDMGGIPREVSQA 469  
DB 410 RAGGLPRHPDALFPPLRLTLFGARYYVLARGLQVEPYPSLQDMGGIPREVSQA 469  
QY 470 LPRPDGSIIFPRDRTWRLDQAKLTQATTSGRWATELPMWGCWHAWSGSLF 520  
DB 470 LPRPDGSIIFPRDRTWRLDQAKLTQATTSGRWATELPMWGCWHAWSGSLF 520

RESULT 12  
ABU66917

ID ABU66917 standard; protein; 520 AA.  
XX  
AC ABU66917;  
XX  
DT 27-MAY-2003 (first entry)  
XX  
DE Human secreted/transmembrane, PRO, protein SEQ ID 144.  
XX  
KW Human; secreted protein; transmembrane protein; PRO;  
KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;  
KW infertility; birth defects; premature aging; AIDS; biosensor;  
KW acquired immunodeficiency syndrome; cancer; diabetic complication;  
KW bioresactor; tumour.  
XX  
OS Homo sapiens.  
XX  
PN US2003032155-A1.  
XX  
PD 13-FEB-2003.  
XX  
PF 03-MAY-2002; 2002US-00137865.  
XX  
PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030999.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005601.  
PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000MO-US005746.  
PR 02-MAR-2000; 2000MO-US005841.  
PR 10-MAR-2000; 2000MO-US006319.  
PR 15-MAR-2000; 2000MO-US006884.  
PR 20-MAR-2000; 2000MO-US007377.  
PR 21-MAR-2000; 2000MO-US007533.  
PR 30-MAR-2000; 2000MO-US008439.  
PR 17-MAY-2000; 2000MO-US013705.  
PR 22-MAY-2000; 2000MO-US014042.  
PR 30-MAY-2000; 2000MO-US014941.  
PR 02-JUN-2000; 2000MO-US015264.  
PR 28-JUN-2000; 2000MO-US020710.  
PR 11-AUG-2000; 2000MO-US020731.  
PR 23-AUG-2000; 2000MO-US023522.  
PR 24-AUG-2000; 2000MO-US023328.  
PR 08-NOV-2000; 2000MO-US030952.  
PR 10-NOV-2000; 2000MO-US030873.  
PR 01-DEC-2000; 2000MO-US032678.  
PR 20-DEC-2000; 2000MO-US034725.  
PR 20-DEC-2000; 2000MO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 01-MAR-2001; 2001MO-US006520.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 18-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00860028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001US-00871092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001MO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001MO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001MO-US020116.  
PR 29-JUN-2001; 2001MO-US021066.  
PR 09-JUL-2001; 2001US-00908827.  
PR 18-JUL-2001; 2001US-00924419.  
PR 06-AUG-2001; 2001US-00927796.  
PR 09-AUG-2001; 2001US-00931836.  
PR 16-AUG-2001; 2001US-00028072.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
FI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
FI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
DR WPI: 2003-331925/31.  
DR N-PSDB; ACA04095.  
XX  
XX New secreted and transmembrane nucleic acids and polypeptides, designated  
PT as PRO, useful for treating inflammation, organ failure, atherosclerosis,  
PT cardiac injury, infertility, birth defects, premature aging, AIDS, or  
PT cancer.  
XX  
PS Claim 12; Fig 144; 659pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising, or which is  
CC at least 80% identical to, or the full-length coding sequence of, any of  
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
CC (one of 275 secreted or transmembrane proteins). The nucleic acid further  
CC comprises the full-length coding sequence of the DNA deposited under  
CC American Type Culture Collection (ATCC) accession number in a list given  
CC in the specification. Also included are vectors and host cells for  
CC producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO

CC extracellular domains and mature sequences, methods of detecting PRO  
CC proteins, methods for stimulating the release of TNF-alpha (tumour  
CC necrosis factor alpha) from human blood, (and the proliferation of  
CC differentiation of chondrocyte cells, the proliferation of, or gene  
CC expression in pericyte cells, the release of proteoglycans from  
CC cartilage, proliferation of inner ear utricular supporting cells, the  
CC proliferation of T-lymphocyte cells, the release of a cytokine from  
CC peripheral blood mononuclear cells (PBMC), or the proliferation of  
CC endothelial cells), a method for modulating the uptake of glucose or free  
CC fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the  
CC binding of A-peptide to factor VIIA, or the differentiation of adipocyte  
CC cells, a method for detecting the presence of a tumour in a mammal and an  
CC oligonucleotide probe derived from any of the nucleotide sequences cited  
CC above. The nucleic acids and polypeptides are useful for treating  
CC inflammatory diseases, organ failure, atherosclerosis, cardiac injury,  
CC infertility, birth defects, premature aging, AIDS (acquired  
CC immunodeficiency syndrome), cancer, or diabetic complications. The  
CC nucleic acids are useful as hybridisation probes, in chromosome and gene  
CC mapping, and in generating antisense RNA or DNA. The polypeptides are  
CC useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both  
CC are useful in tissue typing. The present sequence represents a PRO  
CC protein of the invention  
XX  
SQ Sequence 520 AA:  
Query Match 87.2%; Score 2472; DB 6; Length 520;  
Best Local Similarity 88.9%; Pred. No. 4e-218;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;  
QY 1 MVARVGLLRALQLLLMGHLDAPAEKGGQELREAEAEFLKGYLYNEQVKAPTSTFS 60  
DB 1 MVARVGLLRALQLLLMGHLDAPAEKGGQELREAEAEFLKGYLYNEQVKAPTSTFS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTPRCGVDTNVSYAMARRISDLPRRHTKRRK 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTPRCGVDTNVSYAMARRISDLPRRHTKRRK 120  
QY 121 KRFAKQGNKMYKQHLSTRLVWPEHLRSRQFACAPSSCGATS---QRWS--SGRPOP 175  
DB 121 KRFAKQGNKMYKQHLSTRLVWPEHLRSRQFACAPSSCGATS---QRWS--SGRPOP 175  
QY 176 QAPRTSSPSS-----KGTITMGWAMPMLAQAQAPWTPPLPRRGEAHPQDERWSLSRR 229  
DB 172 EAPAT--GPADIRLTFPGDGHNDLGNAPFGGALAAFLPRRGEAHPQDERWSLSRR 229  
QY 230 RGRNLFVYLAEHIGHTGLTHSPAPRALMAPYKRLGRDALLSMDVLAQVSLYKGPLGG 289  
DB 230 RGRNLFVYLAEHIGHTGLTHSPAPRALMAPYKRLGRDALLSMDVLAQVSLYKGPLGG 289  
QY 290 SVAVOLPGKLFDTDETWDYSPOGRREPOTGPKYCHSSFDAITVDRQOQLYIFKSHFWE 349  
DB 290 SVAVOLPGKLFDTDETWDYSPOGRREPOTGPKYCHSSFDAITVDRQOQLYIFKSHFWE 349  
QY 350 VAADGNVSEPRPQERWVGCPENIEAANVSLNDGDFYFFFGSGRCMRPRGKPVGWLPLQC 409  
DB 350 VAADGNVSEPRPQERWVGCPENIEAANVSLNDGDFYFFFGSGRCMRPRGKPVGWLPLQC 409  
QY 410 RAGGLPRHPDAAFLFPPLRLRLIFKGARVYVLARGGLOVPPYPRSLQDMDGSIPEEVSQA 469  
DB 410 RAGGLPRHPDAAFLFPPLRLRLIFKGARVYVLARGGLOVPPYPRSLQDMDGSIPEEVSQA 469  
QY 470 LPRPDGSIIFRDRRYWRLDQAKLQATTSGMAATELPMWGCWANGSALF 520  
DB 470 LPRPDGSIIFRDRRYWRLDQAKLQATTSGMAATELPMWGCWANGSALF 520  
RESULT 13  
ADA45663  
ID ADA45663 standard; protein; 520 AA.  
XX  
AC ADA45663;  
XX  
DT 20-NOV-2003 (first entry)



XX Novel human secreted and transmembrane protein PRO4339.  
DE Human; secreted and transmembrane protein; PRO;  
XX Tumor necrosis factor alpha release; TNF-alpha release;  
KW Glucose uptake modulator; PFA uptake modulator;  
KW cell proliferation stimulator; cell differentiation stimulator;  
KW cell differentiation inhibitor; cytokine release stimulator; tumor;  
KW lung tumor; colon tumor; breast tumor; prostate tumor; rectal tumor;  
KW cervical tumor; liver tumor; chromosome mapping; gene mapping;  
KW gene therapy; chromosome identification; chromosome marker.  
XX Homo sapiens.  
XX US2003022328-A1.  
PN 30-JAN-2003.  
XX 16-APR-2002; 2002US-00123904.  
XX 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 16-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 05-OCT-1999; 99WO-US021547.  
PR 29-NOV-1999; 99WO-US023089.  
PR 30-NOV-1999; 99WO-US028214.  
PR 01-DEC-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 22-DEC-1999; 99WO-US030999.  
PR 30-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 05-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000WO-US037259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001US-00796520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX (GENTH) GENENTECH INC.  
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerltsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WJ, Zhang Z;  
XX WPI; 2003-584997/55.  
XX N-PSDB; ADA45662.  
XX Novel secreted and transmembrane polypeptide for modulating biological  
PT activity of cell expressing the polypeptide, identifying agonists or  
PT antagonists of polypeptide, and as molecular weight markers.  
XX Claim 12; Fig 144; 659p; English.  
XX The invention describes 305 nucleic acids encoding PRO (secreted and  
CC transmembrane) polypeptides (I). (I) is useful for stimulating the  
CC release of TNF-alpha from human blood, for modulating the uptake of  
CC glucose or PFA by skeletal muscle cells or adipocyte cells, for  
CC stimulating the proliferation or differentiation of chondrocyte cells,  
CC for stimulating the proliferation of or gene expression in pericyte  
CC cells, for stimulating the release of proteoglycans from cartilage, for  
CC stimulating the proliferation of inner ear utricular supporting cells,  
CC for stimulating the proliferation of T-lymphocyte cells, for stimulating  
CC the release of a cytokine from PBMC cells, for inhibiting the binding of  
CC A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte

CC cells, for stimulating proliferation of endothelial cells, for detecting  
CC the presence of tumour in a mammal. The tumour is lung, colon, breast,  
CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes  
CC are useful for isolating genomic and cDNA nucleotide sequences or  
CC antisense probes. (i) is also useful as therapeutic agent. PRO is useful  
CC in assays to identify other proteins or molecules involved in binding  
CC interaction. A polynucleotide (ii) encoding (i) is useful in chromosome  
CC and gene mapping. In generation of antisense RNA and DNA, in the  
CC preparation of PRO polypeptide, for generating transgenic animals or  
CC knockout animals which in turn are useful in the development and  
CC screening of therapeutically useful reagents, in gene therapy, for  
CC chromosome identification, as chromosome marker, and for generating  
CC probes. An anti-(i)-antibody is useful in diagnostic assays for PRO, e.g.  
CC detecting its expression in specific cells, tissues or serum, and for  
CC affinity purification of PRO from recombinant cell culture or natural  
CC sources. (i) and (ii) are useful for tissue typing. This is the amino  
CC acid sequence of a novel human secreted and transmembrane PRO  
CC polypeptide.  
XX

SQ Sequence 520 AA;

Query Match 87.2%; Score 2472; DB 6; Length 520;  
Best Local Similarity 88.9%; Pred. No. 4e-218;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQLLMGHIDAPARGGELRKEAEFLKKGYNQVPAPISTRPS 60  
Db 1 MVARVGLLRALQLLMGHIDAPARGGELRKEAEFLKKGYNQVPAPISTRPS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRATLRQMTPRCGVTDNSYAAMERISDFARHRTKMRX 120  
Db 61 DAIRAFQWVSQLPVSGVLDRATLRQMTPRCGVTDNSYAAMERISDFARHRTKMRX 120  
QY 121 KPRAKGNMXYKHLSYRLVNMPEHRSQFGAPCAPSSCGATS---GRSOP 175  
Db 121 KPRAKGNMXYKHLSYRLVNMPEHRSQFGAPCAPSSCGATS---GRSOP 175  
QY 176 QAPLTSGSPSS-----KGTTWGMAMPMAQGAAPRTRPBRGEAHPDDEKSLRR 229  
Db 176 QAPLTSGSPSS-----KGTTWGMAMPMAQGAAPRTRPBRGEAHPDDEKSLRR 229  
QY 172 EAPAT--GPADIRLTFPGQDNDHDLGNADGPGALANAFLPRGEAHPDDEKSLRR 229  
Db 172 EAPAT--GPADIRLTFPGQDNDHDLGNADGPGALANAFLPRGEAHPDDEKSLRR 229  
QY 230 RGRNLFVYLAHEIGTGLTTHSPAPRALMAPYKXGLGRDALSMQDVLAQSLVCKPLGG 289  
Db 230 RGRNLFVYLAHEIGTGLTTHSPAPRALMAPYKXGLGRDALSMQDVLAQSLVCKPLGG 289  
QY 230 RGRNLFVYLAHEIGTGLTTHSPAPRALMAPYKXGLGRDALSMQDVLAQSLVCKPLGG 289  
Db 230 RGRNLFVYLAHEIGTGLTTHSPAPRALMAPYKXGLGRDALSMQDVLAQSLVCKPLGG 289  
QY 290 SVAVOLPGKLTDFEFTWDSYSPQGRRPETQGPKYCHSSFDALTYDRQQLYIFKGSHEWE 349  
Db 290 SVAVOLPGKLTDFEFTWDSYSPQGRRPETQGPKYCHSSFDALTYDRQQLYIFKGSHEWE 349  
QY 350 VAADGNVSEPRPLQGRWVGLPENIEBAAVSLNDGDFYFFKGGRCMRFRGPKVWGLPQLC 409  
Db 350 VAADGNVSEPRPLQGRWVGLPENIEBAAVSLNDGDFYFFKGGRCMRFRGPKVWGLPQLC 409  
QY 410 RAGGLPRHPDALFPFPLRLILFKGARYVLARGLQVEPYPRSLQDWGIGIEEYSGA 469  
Db 410 RAGGLPRHPDALFPFPLRLILFKGARYVLARGLQVEPYPRSLQDWGIGIEEYSGA 469  
QY 470 LPRPDGSIIFPRDRYRMLDQAKLQATTSGRWATELPMWGMHANSALP 520  
Db 470 LPRPDGSIIFPRDRYRMLDQAKLQATTSGRWATELPMWGMHANSALP 520

RESULT 14

ADA76094 standard; protein, 520 AA.

ADA76094;

20-NOV-2003 (first entry)

Human PRO polypeptide #72.

Human; PRO; secreted polypeptide; transmembrane polypeptide;

KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;  
KW cancer; adrenal; lung; colon; breast; prostate; rectum; kidney; cervix;  
KW liver; microvascular endothelial cell; glucose; FFA;  
KW skeletal muscle cell; adipocyte cell; pericyte cell;  
KW inner ear utricular supporting cell; T-lymphocyte cell;  
KW endothelial cell tube formation; bone disorder; cartilage disorder;  
KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;  
KW rheumatoid arthritis; hemoglobin-associated disorder thalassemia;  
KW immune system cell infiltration.  
OS Homo sapiens.  
PN US2003073212-A1.  
XX 17-APR-2003.  
PD 16-APR-2002; 2002US-00123903.  
XX 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-OCT-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 30-NOV-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030099.  
PR 20-DEC-1999; 99WO-US030911.  
PR 20-DEC-1999; 99WO-US030911.  
PR 22-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 01-MAR-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 10-MAR-2000; 2000WO-US005841.  
PR 15-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.

20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007533.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023523.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00742259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001WO-US002706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 18-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00860208.  
PR 25-MAY-2001; 2001US-00866034.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
PA (GENTH ) GEMENTECH INC.  
XX  
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff B, Gao W;  
PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AU, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
DR MPI, 2003-687639/65.  
DR N-PsDB; ADA76093.  
XX  
PT New isolated nucleic acid encoding a secreted and transmembrane  
PT polypeptide, designated e.g. PRO114 or PRO4978, useful in chromosome and  
PT gene mapping, in generating antisense RNA and DNA, and in gene therapy.  
XX  
PS Claim 12; Fig 144; 659pp; English.  
XX  
XX The invention relates to isolated human PRO polypeptides (secreted and  
XX transmembrane polypeptides) and the polynucleotides encoding them. The  
XX invention also relates to an antibody which specifically binds to a PRO  
XX polypeptide, a method for stimulating the release of tumour necrosis  
XX factor-alpha (TNF-alpha) from human blood, a method for stimulating the  
XX proliferation or differentiation of chondrocyte cells and a method for  
XX detecting the presence of a tumour in a mammal (e.g. adrenal, lung,  
XX colon, breast, prostate, rectal, kidney, cervical and liver tumours). The  
XX polynucleotides are useful in molecular biology, including uses as  
XX hybridisation probes, in chromosome and gene mapping, in generating  
XX antisense RNA and DNA and in gene therapy. The polynucleotides may also  
XX be used in preparing PRO polypeptides by recombinant techniques and in  
XX generating either transgenic animals or knock-out animals which are

CC useful in the development and screening of therapeutically useful  
CC reagents. The PRO polypeptides or antibodies are used in preparing a  
CC medicament for treating a condition responsive to the polypeptides or  
CC antibodies, such as tumours, for stimulating and inhibiting proliferation  
CC of human microvascular endothelial cells, for modulating the uptake of  
CC glucose or FFA by skeletal muscle cells or adipocyte cells, for  
CC stimulating differentiation of adipocyte cells, for stimulating  
CC proliferation of or gene expression in pericyte cells, for stimulating  
CC the proliferation of inner ear utricular supporting cells or T-lymphocyte  
CC cells, for inducing endothelial cell tube formation and for treating  
CC various bone and/or cartilage disorders such as sports injuries and  
CC arthritis. PRO polypeptides which stimulate the release of proteoglycans  
CC from cartilage are useful for treating sports-related joint problems,  
CC articular cartilage defects, osteoarthritis and rheumatoid arthritis. PRO  
CC polypeptides are also useful for treating various mammalian haemoglobin-  
CC associated disorders such as various thalassemias and conditions which  
CC may benefit from enhanced local immune system cell infiltration. This  
CC sequence represents a human PRO polypeptide of the invention. Note: The  
CC sequence data for this patent is also available in electronic format from  
CC USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 520 AA;  
Query Match 87.2%; Score 2472; DB 6; Length 520;  
Best Local Similarity 88.9%; Pred. No. 4e-218;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;  
QY 1 MWARVGLLRALQLLMLKGLDPAERKQERKBAEFLKXYGLVNBQVPAPSTPRS 60  
DB 1 MWARVGLLRALQLLMLKGLDPAERKQERKBAEFLKXYGLVNBQVPAPSTPRS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAIRLROMTPRCQVDTNYSYAAWAERISDLFAHRTYMRKX 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAIRLROMTPRCQVDTNYSYAAWAERISDLFAHRTYMRKX 120  
QY 121 KRFAQGNKMYKQHLSTYLVNMPHLRSRQFCAPCPSSCGATS--GRS--SGRPOP 175  
DB 121 KRFAQGNKMYKQHLSTYLVNMPHLRSRQFCAPCPSSCGATS--GRS--SGRPOP 175  
QY 176 QAPLSSGSSS-----KCTTGMAMPMLMQAPWRPPLPRBEAHPDOERSLSR 229  
DB 172 EAPAT--GPADRLTFQGDNDHDSGNAPFDGFGALAHAFRRKBAHPDOERSLSR 229  
QY 220 RGRNLFVVLAEHIGTLTGITSPAPRALMAPYKRLGSDALISWDVLAQSLYKGPLGG 289  
DB 220 RGRNLFVVLAEHIGTLTGITSPAPRALMAPYKRLGSDALISWDVLAQSLYKGPLGG 289  
QY 230 SVAVOLPGKLFDPFTWDSYSPQGRRPETQGPYCHSSFPDAITVDROQLYIFKGSHEWE 349  
DB 230 SVAVOLPGKLFDPFTWDSYSPQGRRPETQGPYCHSSFPDAITVDROQLYIFKGSHEWE 349  
QY 350 VAAADNVSEPRPLQGRWGLPNTIEAAVSLNDGDFYFFKGRGCRPFGPKRWMLPOLC 409  
DB 350 VAAADNVSEPRPLQGRWGLPNTIEAAVSLNDGDFYFFKGRGCRPFGPKRWMLPOLC 409  
QY 410 RAGGLPRHPDAALFPPLRLILFKGARYVLARGLQVEPYPPLSDWGGIPREVSQA 469  
DB 410 RAGGLPRHPDAALFPPLRLILFKGARYVLARGLQVEPYPPLSDWGGIPREVSQA 469  
QY 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGMHANSGSALF 520  
DB 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGMHANSGSALF 520  
RESULT 15  
ADA18744  
ID ADA18744 standard; protein; 520 AA.  
AC ADA18744;  
XX 20-NOV-2003 (first entry)  
DT Human PRO polypeptide #72.  
XX  
DE

XX Human; PRO; secreted polypeptide; transmembrane polypeptide;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; lung;  
KW colon; breast; prostate; rectum; cervix; liver; tumour; cancer;  
KW glucose uptake; FFA; adipocyte cell; pericyte cell; procollagen;  
KW cartilage; inner ear utricular supporting cell; cytokine; A-peptide;  
KW factor VIIa; endothelial cell.  
XX  
OS Homo sapiens.  
XX  
FN US2003054517-A1.  
XX  
PD 20-MAR-2003.  
XX  
PF 08-MAY-2002; 2002US-00141755.  
XX  
XX 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 16-SEP-1998; 98WO-US019177.  
PR 17-SEP-1998; 98WO-US019330.  
PR 07-OCT-1998; 98WO-US021417.  
PR 29-OCT-1998; 98WO-US022891.  
PR 29-OCT-1998; 98WO-US022892.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028409.  
PR 01-DEC-1999; 99WO-US028301.  
PR 02-DEC-1999; 99WO-US028634.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028564.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 22-DEC-1999; 99WO-US030989.  
PR 30-DEC-1999; 99WO-US030720.  
PR 30-DEC-1999; 99WO-US031243.  
PR 03-JAN-2000; 99WO-US031274.  
PR 06-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 11-FEB-2000; 2000WO-US000376.  
PR 18-FEB-2000; 2000WO-US000365.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004342.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005094.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUN-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00828366.  
PR 28-FEB-2001; 2001WO-US006498.  
PR 01-MAR-2001; 2001WO-US006520.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00860208.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00882636.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
XX (GENTH) GENENTECH INC.  
PA Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z,  
XX WPI, 2003-521854/49.  
DR N-PDB; ADA18743.  
XX  
XX New PRO nucleic acid, useful for preparing a composition for treating  
PT e.g., tumors.  
PT  
PT  
XX  
XX  
PS Claim 12; Fig 144; 660p; English.  
XX  
XX The invention relates to isolated human PRO polypeptides (secreted and  
CC transmembrane polypeptides) and the polynucleotides encoding them. The  
CC invention also relates to an antibody which specifically binds to a PRO  
CC polypeptide, a method for stimulating the release of tumour necrosis  
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the  
CC proliferation or differentiation of chondrocyte cells and a method for  
CC detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,  
CC prostate, rectal, cervical and liver tumours). The polynucleotides are  
CC useful in molecular biology, including uses as hybridisation probes, in  
CC chromosome and gene mapping, in generating antisense RNA and DNA and in  
CC gene therapy. The polynucleotides may also be used in preparing PRO  
CC transgenic animals or knock-out animals which are useful in the  
CC development and screening of therapeutically useful reagents. The PRO  
CC polypeptides or antibodies are used in preparing a medicament for  
CC treating a condition responsive to the polypeptides or antibodies, such

as tumours, for modulating the uptake of glucose or FFA by adipocyte cells, for stimulating the proliferation of or gene expression in pericyte cells, for stimulating the release of proteoglycans from cartilage, for stimulating the proliferation of inner ear utricular supporting cells, for stimulating the release of cytokines from PBMC cells, for inhibiting the binding of A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte cells and for stimulating the proliferation of endothelial cells. This sequence represents a human pro polypeptide of the invention. Note: The sequence data for this patent is also available in electronic format from USPRO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

CC  
XX  
SQ Sequence 520 AA;

Query Match 87.2%; Score 2472; DB 6; Length 520;  
Beat Local Similarity 88.9%; Pred. No. 4e-218;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQLLLMGLDAQPAERGGQELRKEAEAFLEKXYLINEQVPKAPTSTPS 60  
DB 1 MVARVGLLRALQLLLMGLDAQPAERGGQELRKEAEAFLEKXYLINEQVPKAPTSTPS 60  
QY 61 DAIRAFQWTSQLPVSGVLDRAFLRQWTRRCGYTDTNSYAANAERISDLFARHRTKMRK 120  
DB 61 DAIRAFQWTSQLPVSGVLDRAFLRQWTRRCGYTDTNSYAANAERISDLFARHRTKMRK 120  
QY 121 KRFARQGNKWKYKQHLSTYRLVNWPEHLRSRQFAPCAPSSCGATS--ORWS--SGRPQ 175  
DB 121 KRFARQGNKWKYKQHLSTYRLVNWPEHLRSRQFAPCAPSSCGATS--ORWS--SGRPQ 175  
QY 176 QAPLTSGSPSS-----KGTTWGMAMPMAQAPWRTPLPRGEGAHFPODERWSLSRR 229  
DB 172 EAPAT--GPADRLRFPFGQDNDGIGNAFDGPGLAHAFLLPRGEGAHFPODERWSLSRR 229  
QY 230 RGRNLFVYLAHEIGTLGTHSPAPRALMAPYKRLGRDALLSMDVLAVOSLYGKPLCG 289  
DB 230 RGRNLFVYLAHEIGTLGTHSPAPRALMAPYKRLGRDALLSMDVLAVOSLYGKPLCG 289  
QY 290 SYAVOLPGKLFDPFTWDSYSPQGRRPETQGPYCHSFPDAITVDROOQLYTFKGSHPWE 349  
DB 290 SYAVOLPGKLFDPFTWDSYSPQGRRPETQGPYCHSFPDAITVDROOQLYTFKGSHPWE 349  
QY 350 VAADGNVSEBRPLQERWVGLPNIIEAAYSLNDGDFYFFKGRGRCWRFRGPKPVWGLPOLC 409  
DB 350 VAADGNVSEBRPLQERWVGLPNIIEAAYSLNDGDFYFFKGRGRCWRFRGPKPVWGLPOLC 409  
QY 410 RAGGLPRHPDALFPPLRLRLILFKGARYYVLARGLQVEPYPRSLQDMGIPPEVSGA 469  
DB 410 RAGGLPRHPDALFPPLRLRLILFKGARYYVLARGLQVEPYPRSLQDMGIPPEVSGA 469  
QY 470 LPRPDGSIIFPRDDRYWRLDQAKLQATTSGRNATELPMWGCWHAHNSGSLP 520  
DB 470 LPRPDGSIIFPRDDRYWRLDQAKLQATTSGRNATELPMWGCWHAHNSGSLP 520

Search completed: June 13, 2005, 17:58:41  
Job time : 165 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 17:16:46 ; Search time 178 Seconds  
(without alignments)  
1495.962 Million cell updates/sec

Title: US-10-791-980-6  
Perfect score: 2834  
Sequence: 1 MVARVGLLRALQLIMGH.....WATLHPMWCWHSNLSALF 520

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 2472  | 87.2        | 520    | 1     | MM28_HUMAN  |
| 2          | 2073  | 73.1        | 520    | 1     | Q8CGV8      |
| 3          | 2025  | 71.5        | 510    | 2     | Q8CGV8      |
| 4          | 1954  | 68.9        | 496    | 2     | Q8BG29      |
| 5          | 1731  | 61.1        | 393    | 2     | Q9B0G8      |
| 6          | 1167  | 41.2        | 497    | 2     | Q9B0G8      |
| 7          | 654   | 23.1        | 130    | 2     | Q9B0G8      |
| 8          | 588   | 20.7        | 574    | 2     | Q7J2J2      |
| 9          | 570   | 20.1        | 624    | 2     | Q7J2J2      |
| 10         | 559   | 19.7        | 567    | 2     | Q9STK3      |
| 11         | 559   | 19.7        | 570    | 2     | Q9STK3      |
| 12         | 559   | 19.7        | 584    | 2     | Q9STK3      |
| 13         | 559   | 19.7        | 613    | 2     | Q9W122      |
| 14         | 551.5 | 19.5        | 621    | 2     | Q7J2J1      |
| 15         | 548   | 19.3        | 658    | 2     | Q7J2J1      |
| 16         | 546   | 19.3        | 658    | 2     | Q7J2J1      |
| 17         | 541.5 | 19.1        | 546    | 1     | MM19_HUMAN  |
| 18         | 540   | 19.1        | 625    | 2     | Q8B6V6      |
| 19         | 540   | 19.1        | 645    | 1     | MM24_HUMAN  |
| 20         | 536   | 18.9        | 618    | 1     | MM24_HUMAN  |
| 21         | 535.5 | 18.9        | 607    | 2     | Q9B0G8      |
| 22         | 534.5 | 18.9        | 607    | 1     | MM15_HUMAN  |
| 23         | 534   | 18.8        | 669    | 1     | MM15_HUMAN  |
| 24         | 532.5 | 18.8        | 607    | 1     | MM16_MOUSE  |
| 25         | 532   | 18.8        | 618    | 1     | MM24_MOUSE  |
| 26         | 531.5 | 18.8        | 607    | 1     | MM16_MOUSE  |
| 27         | 529   | 18.7        | 582    | 2     | Q9B0G8      |
| 28         | 527.5 | 18.6        | 582    | 2     | Q9B0G8      |
| 29         | 521.5 | 18.4        | 527    | 1     | MM19_MOUSE  |
| 30         | 517.5 | 18.3        | 582    | 1     | MM14_MOUSE  |
| 31         | 517.5 | 18.3        | 582    | 2     | Q8B7X2      |

|    |       |      |     |   |            |                     |
|----|-------|------|-----|---|------------|---------------------|
| 32 | 517.5 | 18.3 | 592 | 2 | Q6W5M7     | Q6W5M7 xenopus lae  |
| 33 | 517   | 18.2 | 657 | 1 | MM15_MOUSE | Q54732 mus musculus |
| 34 | 516.5 | 18.2 | 582 | 2 | Q9XSP0     | Q9XSP0 capra hircu  |
| 35 | 515.5 | 18.2 | 582 | 2 | Q6DFU5     | Q6DFU5 mus musculus |
| 36 | 514.5 | 18.2 | 582 | 1 | MM14_HUMAN | P50281 homo sapien  |
| 37 | 514.5 | 18.2 | 582 | 1 | MM14_MOUSE | P53690 mus musculus |
| 38 | 512.5 | 18.1 | 582 | 1 | MM14_RABIT | Q95220 oryctolagus  |
| 39 | 510.5 | 18.0 | 582 | 2 | Q9GLE4     | Q9GLE4 bos taurus   |
| 40 | 509.5 | 18.0 | 580 | 1 | MM14_PIG   | Q9X590 sus scrofa   |
| 41 | 509   | 18.0 | 613 | 2 | Q919J7     | Q919J7 acylchlorin  |
| 42 | 501   | 17.7 | 431 | 2 | Q9BTF3     | Q9BTF3 oryzae lat   |
| 43 | 497   | 17.5 | 607 | 2 | Q76L07     | Q76L07 oryzae lat   |
| 44 | 496   | 17.5 | 608 | 2 | Q9B947     | Q9B947 gallus gall  |
| 45 | 485   | 17.1 | 562 | 1 | MM25_HUMAN | Q9npa2 homo sapien  |

## ALIGNMENTS

RESULT 1  
MM28\_HUMAN STANDARD; PRT; 520 AA.  
ID MM28\_HUMAN  
AC Q9H239; Q9BTE2; 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Matrix metalloproteinase-28 precursor (EC 3.4.24.-) (MMP-28)  
DN (Epilysin) (UniProtKB/TrEMBL)  
GN Name=MMP28; Synonyms=MMP25;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=21153427; PubMed=11255011; DOI=10.1016/S0378-1119(01)00360-2;  
RT Marchenko G.N., Strongin A.Y.;  
RT "MMP-28, a new human matrix metalloproteinase with an unusual  
RT cysteine-switch sequence is widely expressed in tumors.";  
RT Gene 265:87-93(2001).  
RN [2]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RC TISSUE=Testis;  
RX MEDLINE=21167821; PubMed=11213398; DOI=10.1074/jbc.M001599200;  
RT Lohi J., Wilson C.L., Roby J.D., Parks W.C.;  
RT "Epilysin, a novel human matrix metalloproteinase (MMP-28) expressed  
RT in testis and keratinocytes and in response to injury.";  
RT J. Biol. Chem. 276:10134-10144(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC Southan C., Hughes S.A.;  
RT "Cloning and genomic localization of a novel matrix metalloproteinase.";  
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
RT Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Bush J.,  
RT Chen J., Chow B., Chui C., Crowley C., Curran B., Deuel B., Dowd P.,  
RT Hutton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heidens S.,  
RT Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
RT Lewis L., Liao D., Mark M., Robble E., Sanchez C., Schoenfeld J.,  
RT Seshagiri S., Simons L., Singh J., Smith V., Stinson J., Vagts A.,  
RT Vanden R., Watanabe C., Wiand D., Woods K., Xie M.-H., Yamaura D.,  
RT Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
RA Godowski P., Gray A.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale  
RT effort to identify novel human secreted and transmembrane proteins: a  
RT bioinformatics assessment.";  
RL Genome Res. 13:2265-2270(2003).  
CC -!- FUNCTION: Can degrade casein. Could play a role in tissues  
CC homeostasis and repair.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit, calcium (by similarity).

|                           |  |  |
|---------------------------|--|--|
| CC                        | -1- SUBCELLULAR LOCATION:  | Secreted.  |
| CC                        | -1- ALTERNATIVE PRODUCTS:  | Event=Alternative splicing; Named isoforms=1;                      |
| CC                        |  | Comment=At least 2 isoforms may be produced;                       |
| CC                        | Name=1;  |  |
| CC                        | -1- TISSUE SPECIFICITY:  | IsoId=Q9H239-1; Sequence=displayed;                                |
| CC                        |  | low levels are detected in kidney, pancreas and lung.              |
| CC                        |  | expressed in fetal lung, brain, skeletal muscle and kidney.        |
| CC                        |  | Expressed selectively in keratinocytes. Widely expressed in        |
| CC                        |  | several carcinomas as well. Is up-regulated in response to injury  |
| CC                        |  | in the skin.   |
| CC                        | -1- PTM:   | The precursor is cleaved by a furin endopeptidase (By similarity). |
| CC                        | -1- SIMILARITY:  | Belongs to the peptidase M10A family.                              |
| CC                        | -1- SIMILARITY:  | Contains 1 hemopexin-like domain.                                  |
| CC                        |  |  |
| CC                        | This SWISS-PROT entry is copyright. It is produced through a collaboration   |  |
| CC                        | between the Swiss Institute of Bioinformatics and the EMBL outstation -  |  |
| CC                        | the European Bioinformatics Institute. There are no restrictions on its  |  |
| CC                        | use by non-profit institutions as long as its content is in no way   |  |
| CC                        | modified and this statement is not removed. Usage by and for commercial  |  |
| CC                        | entities requires a license agreement (See <a href="http://www.isdb-sib.ch/announce/">http://www.isdb-sib.ch/announce/</a> |  |
| CC                        | or send an email to <a href="mailto:license@isdb-sib.ch">license@isdb-sib.ch</a> ).  |  |
| CC                        |  |  |
| DR                        | EMBL; AF315683; AACG41981.1; -   |  |
| DR                        | EMBL; AF219624; AAK01480.1; -  |  |
| DR                        | EMBL; AF310002; AAK01706.1; -  |  |
| DR                        | EMBL; AY358987; AAQ89346.1; -  |  |
| DR                        | HSSP; P03956; 1CGU.  |  |
| DR                        | MEROPS; M10_030; -   |  |
| DR                        | Genew; HGNC:14366; MMP28.  |  |
| DR                        | GO; GO:0005578; C:extracellular matrix; NAS.   |  |
| DR                        | GO; GO:0004222; F:metalloendopeptidase activity; TAS.  |  |
| DR                        | GO; GO:0006508; P:proteolysis and peptidolysis; NAS.   |  |
| DR                        | InterPro; IPR000585; Hemopexin.  |  |
| DR                        | InterPro; IPR001818; Pept_M10A_M12B.   |  |
| DR                        | InterPro; IPR006025; Pept_M_Zn_BS.   |  |
| DR                        | InterPro; IPR006026; Peptidase_M.  |  |
| DR                        | InterPro; IPR009070; PGSD_like.  |  |
| DR                        | Pfam; PF00045; Hemopexin; 4.   |  |
| DR                        | Pfam; PF00413; Peptidase_M10; 1.   |  |
| DR                        | Pfam; PF03933; Peptidase_M10_N; 1.   |  |
| DR                        | PRINTS; PR00138; MATRIXIN.   |  |
| DR                        | SMART; SM00120; HX; 4.   |  |
| DR                        | SMART; SM00235; Zmc; 1.  |  |
| DR                        | PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.  |  |
| DR                        | PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.  |  |
| DR                        | PROSITE; PS00142; ZINC_PROTEASE; 1.  |  |
| KW                        | Alternative splicing; Calcium; Extracellular matrix; Glycoprotein;   |  |
| KW                        | Hydrolase; Metalloproteinase; Signal; Zinc; Zymogen.   |  |
| FT                        | SIGNAL   | 1..22 Potential.   |
| FT                        | PROPEP   | 23..122 By similarity.   |
| FT                        | CHAIN  | 123..520 Matrix metalloproteinase-28.                              |
| FT                        | DOMAIN   | 328..520 Hemopexin-like.   |
| FT                        | SITE   | 91..91 Cysteine switch (Potential).                                |
| FT                        | METAL  | 240..240 Zinc (catalytic) (By similarity).                         |
| FT                        | ACT_SITE   | 241..241 By similarity.  |
| FT                        | METAL  | 244..244 Zinc (catalytic) (By similarity).                         |
| FT                        | METAL  | 250..250 Zinc (catalytic) (By similarity).                         |
| FT                        | DISULFID   | 324..510 By similarity.  |
| FT                        | CABOHYD  | 164..164 N-linked (GLNac.) (Potential).                            |
| FT                        | CABOHYD  | 355..355 N-linked (GLNac.) (Potential).                            |
| FT                        | CONFLICT   | 487..487 R -> C (in Ref. 1).                                       |
| SO                        | SEQUENCE   | 520 AA; 58939 MW; BBS7VADA3065B063 CRC64;                          |
| Query Match               |  | 87.2%; Score 2472; DB 1; Length 520;                               |
| Best Local Similarity     |  | 88.9%; Pred. No. 2,2e-177;   |
| Matches 472; Conservative | 4; Mismatches 33; Indels 22; Gaps 5;   |  |
| Oy                        | 1  | MVARVGILLRLATOLLNGHIDAOPAEKGGELRKEAAFLKKGYLNEDVPKAPSTSTRS 60       |
| Dsb                       | 1  | MVARVGILLRLAQLLNGHDPAEKGGELRKEAAFLKKGYLNEDVPKAPSTSTRS 60           |

| QY       | 61   | DARRAOWMSOLPVSGLVDRATROMTTRPCGTTDINSIAAEMRISDLFARRHTYMRK              | 120                       |
|----------|--|---|---------------------------|
| Db       | 61   | DARRAOWMSOLPVSGLVDRATROMTTRPCGTTDINSIAAEMRISDLFARRHTYMRK <td>120</td> | 120                       |
| QY       | 121  | KRFAGKQNMWYKQHLSTYRLVNMPEHLSRQFGAPCAPSPSSCGATS--GRWS--SGRQDP          | 175                       |
| Db       | 121  | KRFAGKQNMWYKQHLSTYRLVNMPEHLSRQFGAPCAPSPSSCGATS--GRWS--SGRQDP          | 175                       |
| QY       | 176  | QALVTSGSYS-----KQTTMGAMMLMAGAPWRTFPLPRGGAHPDODERWLSRR                 | 229                       |
| Db       | 176  | QALVTSGSYS-----KQTTMGAMMLMAGAPWRTFPLPRGGAHPDODERWLSRR                 | 229                       |
| QY       | 230  | RGRNLFVVLAAHEIGTLGLTHSPARALMAPYKRLGRDALSWDVLAVOSLYGKPLG               | 289                       |
| Db       | 230  | RGRNLFVVLAAHEIGTLGLTHSPARALMAPYKRLGRDALSWDVLAVOSLYGKPLG               | 289                       |
| QY       | 290  | SYAVQLPGKLFDFEFWMSYSFQGRRPETQSPRYCHSSFDATITVDRQQLYIFKSGHWE            | 349                       |
| Db       | 290  | SYAVQLPGKLFDFEFWMSYSFQGRRPETQSPRYCHSSFDATITVDRQQLYIFKSGHWE            | 349                       |
| QY       | 350  | VVADGVNSPRLQSEWGLPENIEAAVSLNDGDFYFPKGGRCMRFRPKVWGLPOLC                | 409                       |
| Db       | 350  | VVADGVNSPRLQSEWGLPENIEAAVSLNDGDFYFPKGGRCMRFRPKVWGLPOLC                | 409                       |
| QY       | 410  | RAGGLPRHPDALFPEPLRRLILFKGARYVYLARGLQVEPYPRSLQDWGIPREVSQA              | 469                       |
| Db       | 410  | RAGGLPRHPDALFPEPLRRLILFKGARYVYLARGLQVEPYPRSLQDWGIPREVSQA              | 469                       |
| QY       | 470  | LRPPGSIIFPPDDRYWRDQAKLQATTSGRWATLPMWGCHANGSLFLF                       | 520                       |
| Db       | 470  | LRPPGSIIFPPDDRYWRDQAKLQATTSGRWATLPMWGCHANGSLFLF                       | 520                       |
| RESULT 2 |  |   |                           |
| QY       | 08CGV8   | PRELIMINARY;  | PRT; 520 AA.              |
| Db       | 08CGV8   | PRELIMINARY;  | PRT; 520 AA.              |
| QY       | 01-MAR-2003  | (TREMBLrel. 23, Created)  |                           |
| Db       | 01-MAR-2003  | (TREMBLrel. 23, Last sequence update)                                 |                           |
| QY       | 01-MAR-2004  | (TREMBLrel. 26, Last annotation update)                               |                           |
| Db       | 01-MAR-2004  | (TREMBLrel. 26, Last annotation update)                               |                           |
| QY       | Matrix   | metalloprotease-28 precursor.   |                           |
| Db       | Matrix   | metalloprotease-28 precursor.   |                           |
| QY       | NCBI   | Textid=10090;   |                           |
| Db       | NCBI   | Textid=10090;   |                           |
| QY       | SEQUENCE FROM N.A.   |   |                           |
| Db       | SEQUENCE FROM N.A.   |   |                           |
| QY       | STRAIN=BALB/c; TISSUE=Heart;                                 |   |                           |
| Db       | STRAIN=BALB/c; TISSUE=Heart;                                 |   |                           |
| QY       | Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.      |   |                           |
| Db       | Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.      |   |                           |
| QY       | EMBL; AY065653; AAL47576.1; -                                |   |                           |
| Db       | EMBL; AY065653; AAL47576.1; -                                |   |                           |
| QY       | HSSP; P03956; 1CGL   |   |                           |
| Db       | HSSP; P03956; 1CGL   |   |                           |
| QY       | MEROPS; M10.030; -   |   |                           |
| Db       | MEROPS; M10.030; -   |   |                           |
| QY       | GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA. |   |                           |
| Db       | GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA. |   |                           |
| QY       | GO; GO:0004222; F:metalloendopeptidase activity; IEA.        |   |                           |
| Db       | GO; GO:0004222; F:metalloendopeptidase activity; IEA.        |   |                           |
| QY       | InterPro; IPR000508; P:proteolysis and peptidolysis; IEA.    |   |                           |
| Db       | InterPro; IPR000508; P:proteolysis and peptidolysis; IEA.    |   |                           |
| QY       | InterPro; IPR006026; Peptidase_M.                            |   |                           |
| Db       | InterPro; IPR006026; Peptidase_M.                            |   |                           |
| QY       | InterPro; IPR001818; Pept_M10A_M12B.                         |   |                           |
| Db       | InterPro; IPR001818; Pept_M10A_M12B.                         |   |                           |
| QY       | InterPro; IPR009070; PGSD_Like.                              |   |                           |
| Db       | InterPro; IPR009070; PGSD_Like.                              |   |                           |
| QY       | Pfam; PF00045; Hemopexin; 4.                                 |   |                           |
| Db       | Pfam; PF00045; Hemopexin; 4.                                 |   |                           |
| QY       | PRINTS; PRO0138; MATRIKIN.                                   |   |                           |
| Db       | PRINTS; PRO0138; MATRIKIN.                                   |   |                           |
| QY       | SMART; SMO0120; HX; 4.                                       |   |                           |
| Db       | SMART; SMO0120; HX; 4.                                       |   |                           |
| QY       | SMART; SMO0235; ZHMC; 1.                                     |   |                           |
| Db       | SMART; SMO0235; ZHMC; 1.                                     |   |                           |
| QY       | PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.                  |   |                           |
| Db       | PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.                  |   |                           |
| QY       | SIGNAL.  |   |                           |
| Db       | SIGNAL.  |   |                           |
| QY       | 123  | 520   | matrix metalloprotease-28 |
| Db       | 123  | 520   | matrix metalloprotease-28 |



SQ SEQUENCE 520 AA; 58664 MW; 69924501CFC1991D CRC64;

Query Match 73.1%; Score 2073; DB 2; Length 520;  
 Best Local Similarity 74.4%; Pred. No. 2, 2e-147;  
 Matches 395; Conservative 33; Mismatches 81; Indels 22; Gaps 5;

QY 1 MVARVLLRLALQLLLMGLHDAQPAERGQGLRKEAEAFLEKYGYNQVPAKPTSTRS 60  
 DB 1 MVAGVLLRLALPLLLMGCDQAQPTQHGRLPELRQEAFALEKYGYSIQGSKAPASQPR 60  
 QY 61 DAIRAFQWVSQPLVSGVLDRAATLRQMTTRPCGVTDTNSYAAMERISDLFARHTQMRK 120  
 DB 61 NAIRFQWISQPLVSGVLDQATLRQMTTRPCGVAIDTSHATWTERISTLLAGRAQMRK 120  
 QY 121 KRPFAKGNKMYKQHLSTRLVNMPEHLRSRQFGAPCAPSPSCGATS---QRWS---SGRPOP 175  
 DB 121 KRPFAKGNKMYKQHLSTRLVNMPEHLRSRQFGAPCAPSPSCGATS---QRWS---SGRPOP 175  
 QY 176 QAPLTSGPSSS-----KGTTHGMAMPMLAOGAPWRTPPLPRRGAHFDQDERWLSLR 229  
 DB 172 EAPAT--GPADIRLTFQGDHNDGLANAFDQPGALAAHFLPRRGAHFDQDERWLSLR 229  
 QY 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYGRPLG 289  
 DB 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYGRPLG 289  
 QY 290 SVAVALPGKLPFDFTWDSYSPQGRRPETQGRPKYCHSSFDATVDRQOQLYIFKSGHFW 349  
 DB 290 SVAVALPGKLPFDFTWDSYSPQGRRPETQGRPKYCHSSFDATVDRQOQLYIFKSGHFW 349  
 QY 350 VAADGNVSEPRPLQGRWVGLPRNIEAAVSLNDGDFYFFKGRGCMRFRGPKPWGLPOLC 409  
 DB 350 VYVDGNVSEPRPLQGRWVGLPRNIEAAVSLNDGDFYFFKGRGCMRFRGPKPWGLPOLC 409  
 QY 410 RAGGLPRHPDALFPFPLRLILFKGARYYVLARGLQVEPPYPRSLQDMGGLPEEVSQA 469  
 DB 410 RAGGLPRHPDALFPFPLRLILFKGARYYVLARGLQVEPPYPRSLQDMGGLPEEVSQA 469  
 QY 470 LRPDGSIIFFRDHRYWRLDQAKLQATTSGRWATELPMWGCWANSGLALF 520  
 DB 470 LRPDGSIIFFRDHRYWRLDQAKLQATTSGRWATELPMWGCWANSGLALF 520

RESULT 3  
 O8BGV5 PRELIMINARY; PRT; 510 AA.

AC O8BGV5; 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Matrix metalloproteinase-28 variant A precursor.  
 GN Name=Mmp28;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Heart;  
 RA Ilman S.A., Keski-Oja J., Lohi J.;  
 DT Submitted (DEC-2001) to the EMBL/GenBank/DBJ database.  
 DR EMBL; AY071828; AA157761.1; -  
 DR HSSP; P03956; 1CGL.  
 DR MEROPS; M10.010; -  
 DR MGD; MG1:2153062; Mmp28.  
 DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.  
 DR GO; GO:0004222; P:metalloendopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR006058; Hemoxen.  
 DR InterPro; IPR006026; Peptidase\_M.  
 DR InterPro; IPR001818; Pept\_M10A\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR009070; PGBD\_like.

DR Pfam; PF00045; Hemoxen; 4.  
 DR Pfam; PF03933; Peptidase\_M10\_N; 1.  
 DR PRINTS; PR00138; MATRININ.  
 DR SMART; SM00120; HK; 4.  
 DR SMART; SM00235; ZMOC; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW Signal.  
 FT SIGNAL.  
 FT CHAIN 123 510 Potential.  
 FT matrix metalloproteinase-28 variant A precursor.

SQ SEQUENCE 510 AA; 57371 MW; 3479A65940835E93 CRC64;

Query Match 71.5%; Score 2025; DB 2; Length 510;  
 Best Local Similarity 73.3%; Pred. No. 8, 6e-144;  
 Matches 389; Conservative 31; Mismatches 79; Indels 32; Gaps 6;

QY 1 MVARVLLRLALQLLLMGLHDAQPAERGQGLRKEAEAFLEKYGYNQVPAKPTSTRS 60  
 DB 1 MVAGVLLRLALPLLLMGCDQAQPTQHGRLPELRQEAFALEKYGYSIQGSKAPASQPR 60  
 QY 61 DAIRAFQWVSQPLVSGVLDRAATLRQMTTRPCGVTDTNSYAAMERISDLFARHTQMRK 120  
 DB 61 NAIRFQWISQPLVSGVLDQATLRQMTTRPCGVAIDTSHATWTERISTLLAGRAQMRK 120  
 QY 121 KRPFAKGNKMYKQHLSTRLVNMPEHLRSRQFGAPCAPSPSCGATS---QRWS---SGRPOP 175  
 DB 121 KRPFAKGNKMYKQHLSTRLVNMPEHLRSRQFGAPCAPSPSCGATS---QRWS---SGRPOP 175  
 QY 176 QAPLTSGPSSS-----KGTTHGMAMPMLAOGAPWRTPPLPRRGAHFDQDERWLSLR 229  
 DB 172 EAPAT--GPADIRLTFQGDHNDGLANAFDQPGALAAHFLPRRGAHFDQDERWLSLR 229  
 QY 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYGRPLG 289  
 DB 230 RGRNLFVYLAHEIGHTLGLTHSPAPRALMAPYKRLGRDALLSMDVLAQSLYGRPLG 289  
 QY 290 SVAVALPGKLPFDFTWDSYSPQGRRPETQGRPKYCHSSFDATVDRQOQLYIFKSGHFW 349  
 DB 290 SVAVALPGKLPFDFTWDSYSPQGRRPETQGRPKYCHSSFDATVDRQOQLYIFKSGHFW 349  
 QY 350 VAADGNVSEPRPLQGRWVGLPRNIEAAVSLNDGDFYFFKGRGCMRFRGPKPWGLPOLC 409  
 DB 350 VYVDGNVSEPRPLQGRWVGLPRNIEAAVSLNDGDFYFFKGRGCMRFRGPKPWGLPOLC 409  
 QY 410 RAGGLPRHPDALFPFPLRLILFKGARYYVLARGLQVEPPYPRSLQDMGGLPEEVSQA 469  
 DB 410 RAGGLPRHPDALFPFPLRLILFKGARYYVLARGLQVEPPYPRSLQDMGGLPEEVSQA 469  
 QY 470 LRPDGSIIFFRDHRYWRLDQAKLQATTSGRWATELPMWGCWANSGLALF 520  
 DB 460 LRPDGSIIFFRDHRYWRLDQAKLQATTSGRWATELPMWGCWANSGLALF 510

RESULT 4  
 O8BG29 PRELIMINARY; PRT; 496 AA.

AC O8BG29; 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
 DE Matrix metalloproteinase-28 variant B precursor (Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone: D130023p12 product: matrix metalloproteinase 28 (epilysin), full insert sequence).  
 GN Name=Mmp28;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Heart;  
 RA Ilman S.A., Keski-Oja J., Lohi J.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RP [2]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RT Carninci P., Hayashizaki Y.;  
 RL "High-efficiency full-length cDNA cloning.";  
 RM Meth. Enzymol. 303:19-44(1999).  
 RP [3]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RT RIKEN FANTOM Consortium;  
 RL "Functional annotation of a full-length mouse cDNA collection.";  
 RM Nature 409:685-690(2001).  
 RP [4]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RM Nature 420:563-573(2002).  
 RP [5]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RT Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RL "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RM Genome Res. 10:1617-1630(2000).  
 RP [6]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RT Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Kitanaka T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashigagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RL "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";  
 RM Genome Res. 10:1757-1771(2000).  
 RP [7]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Spinal ganglion;  
 RA Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,  
 RA Fukuda S., Futuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaubara T.,  
 RA Kotoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akchira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RM EMBL: AV071829; AAL57762.1; -;  
 DR EMBL: AK051258; BAC34577.1; -;  
 DR HSSP: P03956; 1CGL.  
 DR MEROPS: M10.030; -;  
 DR MGD: MGJ:2153062; Mmp28.  
 DR GO: GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.  
 DR GO: GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO: GO:0005508; F:proteolysis and peptidolysis; IEA.  
 DR InterPro: IPR000585; Hemopexin.  
 DR InterPro: IPR006026; Peptidase M.  
 DR InterPro: IPR001818; Pept\_M0A\_M12B.  
 DR InterPro: IPR006025; Pept\_M\_zn\_BS.

DR InterPro: IPR009070; PGBD\_11ke.  
 DR Pfam: PF00045; Hemopexin\_3.  
 DR PRINTS: PRO03933; Peptidase\_M10\_N; 1.  
 DR SMART: SM00138; MATRIXIN.  
 DR SMART: SM00120; HX; 3.  
 DR SMART: SM00235; Zmc; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KM SIGNAL.  
 FT SIGNAL.  
 FT CHAIN 123 496 Potential.  
 FT matrix metalloproteinase-28 variant B  
 FT precursor.  
 SQ SEQUENCE 496 AA; 55829 MW; CEE16CD0F0DD2906D CRC64;  
 Query Match 68.9%; Score 1954; DB 2; Length 496;  
 Best local similarity 71.2%; Pred. No. 1.8e-138;  
 Matches 378; Conservative 30; Mismatches 77; Indels 46; Gaps 6;  
 QY 1 MYARVGLLRALQLLWGHLDQAPAEKGOELRKAENAFLEKGYLNEQVAPKPTSTFS 60  
 DB 1 MYAGVSLLRALPLLLWGCODAPQTHGLPELRGEAEFLKGYLSQSKAPASQFR 60  
 QY 61 DAIRAFQVQSGLPVSGVLDRAITLQMTPRRCGVDTNYSYAMARISDLFARHRTKRRK 120  
 DB 61 NAIRFQWISQPLPSGVLDQATLRQMTPRRCGVADTSHATWTERISTLLAGHRAKRRK 120  
 QY 121 KRFAQGNKMYKQHLSTYLVMPPEHLRSRQFACAPSSCGATS---QRMS--SGRPOP 175  
 DB 121 KRFAKPGKMYKQHLSTYLVMPPEHLRSRQFACAPSSCGATS---QRMS--SGRPOP 175  
 QY 121 KRFAKPGKMYKQHLSTYLVMPPEHLRSRQFACAPSSCGATS---QRMS--SGRPOP 175  
 DB 121 KRFAKPGKMYKQHLSTYLVMPPEHLRSRQFACAPSSCGATS---QRMS--SGRPOP 175  
 QY 176 QAPLTSGSPSS-----KGTTWGMAMPLOAGAPWPTPLPRRGAHPQDERWSLSRR 229  
 DB 172 BAPPT--GPADIRLTFQGGHNDGIANAFDGGALAAFLPRRGAHPQDERWSLSRR 229  
 QY 230 RGRNLFVYLAHEIGHTGLTSPAPRALMAPYRKLGRDLGLSDVYLAQSLYKRLCG 289  
 DB 230 RGRNLFVYLAHEIGHTGLTSPAPRALMAPYRKLGRDLGLSDVYLAQSLYKRLCG 289  
 QY 290 SVAQVLPKGLFTDEFTWDSYSPQGRPRETQPKYCHSSFPDATTVDROQLYIKGSHFWE 349  
 DB 290 SVAQVLPKGLFTDEFTWDSYSPQGRPRETQPKYCHSSFPDATTVDROQLYIKGSHFWE 349  
 QY 350 VAADGVNVEPRPLQERVGVLPNIEAAVSLNCDPTFFRGKCMWRPRKPYWGLPOLC 409  
 DB 334 -----EPRPQKRPKPGIPGIEAAVSLNCDPTFFRGKCMWRPRKPYWGLPOLC 409  
 QY 410 RAGGLPRHPDALFFPPLRLRLIFKGRYYVYLAARGIQVPEYPRSLQDQGIPEEYSGA 469  
 DB 386 RAGGLPRHPDALFFPPLRLRLVLFKGRYYVYLAQGMQVEPYPRSLRDMAGVPEEYSGA 445  
 QY 470 LPRPDGSIIFPRDRYWRDLQATLQATTSGRWATELPMWGMHANSALF 520  
 DB 446 LPRPDGSIIFPRDRYWRDLQATLQATTSGRWATELPMWGMHANSALF 496  
 RESULT 5  
 ID Q9BUG8 PRELIMINARY; PRT; 393 AA.  
 AC Q9BUG8;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Matrix metalloproteinase 28, preproprotein isoform 2.  
 GN Name:MMP28;  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Caessant T.L., Schetz T.E.,  
 RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RN Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC002631; AA02631.1; -.  
 DR HSSP; P03956; 1CGL.  
 DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000585; Hemopeptin.  
 DR InterPro; IPR006026; Peptidase\_M.  
 DR InterPro; IPR001818; Pept\_M10A\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR009070; RGD\_Like.  
 DR Pfam; PF00065; Hemopeptin; 1.  
 DR Pfam; PF03953; Peptidase\_M10\_N; 1.  
 DR PRINTS; PR00138; MATRXIN.  
 DR SMART; SM00235; ZMNC; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN 1.  
 SO SEQUENCE 393 AA; 4469 MW; 07D443B5401699 CRC64;

Query Match 61.1%; Score 1731; DB 2; Length 393;  
 Best Local Similarity 85.2%; Pred. No. 7.9e-122;  
 Matches 341; Conservative 4; Mismatches 33; Indels 22; Gaps 5;  
 QY 1 MVARVGLLLRAIQLLLMGMHDAQPARSGOELKBAEPLKXGYLNEYPKAPSTRS 60  
 DB 1 MVARVGLLLRAIQLLLMGMHDAQPARSGOELKBAEPLKXGYLNEYPKAPSTRS 60  
 QY 61 DAIRAFQWVSQLPVSGVLDRAITRQMTTRPCGVTDNTSYAAMERISDLFARHTQMRK 120  
 DB 61 DAIRAFQWVSQLPVSGVLDRAITRQMTTRPCGVTDNTSYAAMERISDLFARHTQMRK 120  
 QY 121 KRPFAKGNKMYKOHLSYRLVNPBEHLRSRQFGAPCAPSSCGATS---SGRPOP 175  
 DB 121 KRPFAKGNKMYKOHLSYRLVNPBEHLRSRQFGAPCAPSSCGATS---SGRPOP 175  
 QY 121 KRPFAKGNKMYKOHLSYRLVNPBEHLRSRQFGAPCAPSSCGATS---SGRPOP 175  
 DB 121 KRPFAKGNKMYKOHLSYRLVNPBEHLRSRQFGAPCAPSSCGATS---SGRPOP 175  
 QY 176 QAPLTGSGSPS-----KGTTMGMAMPPLMAOGAPARTPTLPARGENHPQDERKSLSR 229  
 DB 176 QAPLTGSGSPS-----KGTTMGMAMPPLMAOGAPARTPTLPARGENHPQDERKSLSR 229  
 QY 230 RGNLFLVLAHEIGTTLGLTHSPAPRALMAPYKRLGRDALLSMDDVLAQVSLYKGPLG 289  
 DB 230 RGNLFLVLAHEIGTTLGLTHSPAPRALMAPYKRLGRDALLSMDDVLAQVSLYKGPLG 289  
 QY 230 RGNLFLVLAHEIGTTLGLTHSPAPRALMAPYKRLGRDALLSMDDVLAQVSLYKGPLG 289  
 DB 230 RGNLFLVLAHEIGTTLGLTHSPAPRALMAPYKRLGRDALLSMDDVLAQVSLYKGPLG 289  
 QY 290 SVAVQLPGKLTDFETWDSYSPQGRRPETGPRYCHSSPDALTVDROQLYIFKSHFWE 349  
 DB 290 SVAVQLPGKLTDFETWDSYSPQGRRPETGPRYCHSSPDALTVDROQLYIFKSHFWE 349  
 QY 290 SVAVQLPGKLTDFETWDSYSPQGRRPETGPRYCHSSPDALTVDROQLYIFKSHFWE 349  
 DB 290 SVAVQLPGKLTDFETWDSYSPQGRRPETGPRYCHSSPDALTVDROQLYIFKSHFWE 349  
 QY 350 VAADGNVSEPRPLQERWVGLPNIIEAAVSLNDGDFEYFFK 389  
 DB 350 VAADGNVSEPRPLQERWVGLPNIIEAAVSLNDGDFEYFFK 389

RESULT 6  
 ID 06P714 PRELIMINARY; PRT; 497 AA.  
 AC 06P714  
 DT 05-JUL-2004 (TEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)  
 DE MGC68506 protein.  
 GN Name=MGC68506;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodidae; Xenopus.  
 NC NCB1\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Hsieh F.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Caessant T.L., Schetz T.E.,  
 RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RL Klein S., Strausberg R.;  
 RX Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC061659; AA061659.1; -.  
 DR HSSP; P08254; 1B3D.  
 DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000585; Hemopeptin.  
 DR InterPro; IPR006026; Peptidase\_M.  
 DR InterPro; IPR001818; Pept\_M10A\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR009070; RGD\_Like.  
 DR Pfam; PF00065; Hemopeptin; 4.  
 DR Pfam; PF03953; Peptidase\_M10\_N; 1.  
 DR PRINTS; PR00138; MATRXIN.  
 DR SMART; SM00120; HK; 4.  
 DR SMART; SM00235; ZMNC; 1.  
 DR PROSITE; PS00024; HEMOPEPTIN; UNKNOWN 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN 1.  
 SO SEQUENCE 497 AA; 5706 MW; 3D6BA00FB3BD205 CRC64;

Query Match 41.2%; Score 1167; DB 2; Length 497;  
 Best Local Similarity 47.8%; Pred. No. 2.7e-79;



| Query Match           | 20.1%   | Score 570;  | DB 2;       | Length 624; |
|-----------------------|---|---|-------------|-------------|
| Best Local Similarity | 29.6%   | Pred. No. 2,7e-34;  |             |             |
| Matches 164;          | Conservative 85;  | Mismatches 185;   | Indels 120; | Gaps 23;    |
| Qy                    | 9   | LRAQLLHGLHLDQAPRERGQGLKAKAEAFLEKXYGLINEVPRKAPSTRSSD-----A   | 62          |             |
| Db                    | 63  | IRALAIYLVGCAAGTA--SPVSTTPPAELVLSQGLSPRY-TNPTSGLNLDDTWEKA    | 119         |             |
| Qy                    | 63  | IRAFOWSOLPVSGLVRLRATLRONTPRCGTPTDINSAAWABRISDFARHRTQRRKR    | 122         |             |
| Db                    | 120   | IMEVQSAPGLAVTBELDGETMQLMSLRRCVYKDKVG-----SD-----TSSKR       | 164         |             |
| Qy                    | 123   | PAKQNKMYKXOHLSYRLVWPEHLR-----SRQGAPCAPSSCGATSQRWSSGRPQ      | 174         |             |
| Db                    | 165   | YALGSSMWKVDLYRISKYPRRLERTAVDKELAKAGV-----WSEYTDL            | 210         |             |
| Qy                    | 175   | POAPLTSGS-----PSSKGTITMGAMFLMAQAPMTRP-----FLPRGGAHF         | 218         |             |
| Db                    | 211   | RFTPKTKGAVHIDIRFEBENHG-----DSDPFQPGGTLAHAYFPVYGDAHF         | 258         |             |
| Qy                    | 219   | DQDERWLSRRRRGNLFPVLAHEIGTILGTHSPAPALMAFYKRLGRDAL--LSMDV     | 276         |             |
| Db                    | 259   | DDAEQMTIDKRGNLQVAAHGHGSLGSHSVRSALMAFPKR--GYDPVFLDSDDI       | 316         |             |
| Qy                    | 277   | LAVOSLTVG---KPLGGSVAVVLFGKLFPTETWDSYSPQGRRETOGPXYCHS-SFDAI  | 331         |             |
| Db                    | 317   | QGITLTIVGTRNPGGAGAT-----PRTPRPKPTEMDSBLCTSPKIDAI            | 363         |             |
| Qy                    | 332   | TVDRQOQLYIFKSGHFMEVAAADGVSE--PRPLQERVAVGLPUNIEAAVSLNDGPFYFK | 389         |             |
| Db                    | 364   | FNTADGSTYAFKGGKYKL--TENAVAGYEPKSIDGMPGLPGNID--AAFYKNGKTYFFQ | 421         |             |
| Qy                    | 390   | GGRCWRMRGKRPVWGLPQLCRAG--GLPRHPDALFPPRLRLLEFKGARY--VLARG    | 444         |             |
| Db                    | 422   | GTKWRYQAGRTIDGDYKEISEGTTGVDPDHLDAAMVGVGNGKIYFYTGSKWRPDLKRP  | 481         |             |
| Qy                    | 445   | GLQVEPYPSLQDWGIPREVSAGLPRPGSIIFFDDRYWRLDQAKLQATTS-----      | 498         |             |
| Db                    | 482   | --PKSYTPKPISMWGVNSVDAAQLYTTNGYTYFFPDCKYRNDNTFFVVDGSDPPFPR   | 539         |             |
| Qy                    | 499   | --GRWATELPMWGC  | 510         |             |
| Db                    | 540   | PTAHM-----WYGC  | 548         |             |
| RESULT 10             |   |   |             |             |
| Q9GTK3                | PRELIMINARY;  | PRT;  | 567 AA.     |             |
| AC                    | Q9GTK3;   |   |             |             |
| DT                    | 01-MAR-2001 (Tremblurel. 16, Created)                         |   |             |             |
| DT                    | 01-MAR-2001 (Tremblurel. 16, Last sequence update)            |   |             |             |
| DT                    | 01-MAR-2004 (Tremblurel. 26, Last annotation update)          |   |             |             |
| DE                    | Matrix metalloproteinase 1.                                   |   |             |             |
| GN                    | Name=Mmp1;  |   |             |             |
| OS                    | Drosophila melanogaster (fruit fly).                          |   |             |             |
| OC                    | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; |   |             |             |
| CC                    | Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;    |   |             |             |
| CC                    | Ephydroidea; Drosophilidae; Drosophila.                       |   |             |             |
| NCBI                  | taxid=7227;   |   |             |             |
| LN                    | [1]   |   |             |             |

|                           |        |                    |            |             |
|---------------------------|--------|--------------------|------------|-------------|
| Query Match               | 19.7%; | Score 559;         | DB 2;      | Length 567; |
| Best Local Similarity     | 30.4%; | Pred. No. 1.6e-33; |            |             |
| Matches 158; Conservative | 74;    | Mismatches 200;    | Indels 88; | Gaps 10;    |

```

QY 35 EAAFAFEKGYLNEQVPKAPSTIRSD-----AIAAFQWVSLPVSGVLDRALTRQMITR 88
Db 58 QAEIYISQGYLPASA-RNPASSGLHQRTWVAIEEFOFAGLINTGELDEITKMLSL 116
QY 89 PRGCVTDTNVSAAAMERISDLFAHRTMERKKRFAQOGKWKYKOHLSYRLNMPDEHRS 148
Db 117 PRGCVAD-----FVGTGDHSKRYVALQSRMKVKKULTYIKIKYKRLKR 160
QY 149 RQFGAPCAPISSCGATISQWSSGAPQOAPLTGS-----SPSSKGTITWGMAMPL 198
Db 161 VDVA-----EIGAFAVWSEDTDLTFTRKTGPAVHIEIKFEVSEHGDAFDB----- 209
QY 199 MAOGAPWRTPLP-RGEAHFDODERWLSRRRGRLFVLAHEIGHTLGLTHSPAPAL 257
Db 210 --QGGLTAAFPVPGDAHFDDAEIWTISPGTGLFVAAHGHSHGSHSDQSAL 267
QY 258 MAPYKRLGRDALLSMVDVLAQSLYKPKLGGSSVAVQLPGLFTDFETWDSYPOGRAP 317
Db 268 MAPFGRGEPVFKLDEDDKKAIOSLYRKTN-----QLPPTVWYPAITORPYS- 316
QY 318 TGGPK-----YC-HSSEDAITVDROQOLTFKGSHFWEVADGNVSE--PRPLOERNVG 368
Db 317 ---PKVPLDSDICDSKVDTLFNSAQEYTAFFGDKYTLCTLD--SVEGTYQLISKMPG 372
QY 369 LPENLEAAVSLNGDPTFFKGGCHMFRGKPFWMLPOLCRAG--GLPRHPDAALFPP 426
Db 373 LPGNLD-AAFTYKNGKTYFFKGTQYWRKYQGRQMDGVYPEKISEGTTGPDILDAWVWG 431
QY 432 LRLTLFFKARYVYLARGGL-QVEYPRSLQDMGSIPEVSGALPRDGSIIIFRDRY 485
Db 432 NGKTIYFFGSGFWMFPDAKRPVYKATYKPSISWBEVPPNLDALAKTYNGTYTFFKDKY 491
QY 486 WRLDQAKIQAAT-----SGRWATELPMWGCMAHNSGS 517
Db 492 YRFHDARPAVDSATPPPRPTAJM-----WFGKTYPSST 526

```

|              |        |
|--------------|--------|
| RESULT 11    |        |
| Q8MLN6       |        |
| ID           | Q8MLN6 |
| PRELIMINARY; | PRT;   |
|              | 570 AA |

AC OEWLNK;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE CQ4859-PA.  
GN Name=Mpmp; ORFNames=CQ4859;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7227;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blaise R.G., Champe M., Pfeiffer B.D.,  
RA Abail J.F., Adayant A.A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman J.P., Bhandari D., Bolintsov S.,  
RA Borhova D., Botchan M.R., Bouck J., Brooksstein P., Brotlier P.,  
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Casway S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegawa C.,  
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moutreli A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclel J.M.,  
RA Palczak M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reiter K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassenaar D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.W., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri U.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RL "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
[2]  
SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
RA Patel S.A., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Paclel J.M., Pak S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svitskas R., Tabot P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT melanogaster euchromatidic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
[3]  
SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminler J.S., Bergman C.M., Krommiller B., Carlson J., Svitskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomic perspective.";

```

RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Mlera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kunkler J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu U., Berman B.P.,
RA Beltencourt B.R., Celniker S.B., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
RP SEQUENCE FROM N.A.
RG FlyBase;
RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003464; AAM68327.1; -.
DR HSSP; P03956; ICGL.
DR MEROPS; M10.031; -.
DR FlyBase; FBgn0035049; Mmp1.
DR GO; GO:0044222; P:metalloendopeptidase activity; TAS.
DR GO; GO:0048102; P:autophagic cell death; IEP.
DR GO; GO:0035001; P:dorsal trunk growth; IMP.
DR GO; GO:0002168; P:larval development (sensu Insecta); IMP.
DR GO; GO:0035071; P:salivary gland cell death; IEP.
DR InterPro; IPR0060585; Hemopexin.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR009070; PGBD_like.
DR Pfam; PF00045; Hemopexin; 4.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
DR SEQUENCE 570 AA; 63484 MW; 902DCF91974CF0BD CRC64;
SQ
Query Match 19.7%; Score 559; DB 2; Length 570;
Best Local Similarity 30.4%; Pred. No. 1.6e-33;
Matches 158; Conservative 74; Mismatches 200; Indels 88; Gaps 18;

```

```

DB 376 LFGNID-AAFTYKNGKTYFFKGTQYMRYOGRQMDGVYPKESISEFTGIPHLIDAAWYWG 434
QY 427 LRRLILFKGARVYTLARGL-OVEPPYRSLQDMGSIPEEVSGLPRPDGSIIFRDRY 485
DB 435 NGKTYFFKSGFKWFPDPAKRPVVAASYPKTISNMEGVNNIDALKTNGTYTFFKGDY 494
QY 486 WRLLDQAKLQATT-----SGRWATLPMWGCHWANGS 517
DB 495 YRFIDARFAVDSATPPPRPTAH-----WFGCKNTFSST 529

RESULT 12
Q8MT35
ID Q8MT35 PRELIMINARY; PRT; 584 AA.
AC Q8MT35;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE R862222P.
GN Name=Mmp1;
OS Drosophila melanogaster (fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely.
RA Stapleton M., Brokstein P., Hong L., Abdayani A., Carlson J.,
RA Champ M., Chavez C., Dorett V., Dresnek D., Fafan D., Faise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118405; AAM48434.1; -.
DR HSSP; P03956; ICGL.
DR MEROPS; M10.031; -.
DR FlyBase; FBgn0035049; Mmp1.
DR GO; GO:0044222; P:metalloendopeptidase activity; TAS.
DR GO; GO:0048102; P:autophagic cell death; IEP.
DR GO; GO:0035001; P:dorsal trunk growth; IMP.
DR GO; GO:0002168; P:larval development (sensu Insecta); IMP.
DR GO; GO:0035071; P:salivary gland cell death; IEP.
DR InterPro; IPR0060585; Hemopexin.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR009070; PGBD_like.
DR Pfam; PF00045; Hemopexin; 4.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
DR SEQUENCE 584 AA; 64670 MW; F903AED21F9187B CRC64;
SQ
Query Match 19.7%; Score 559; DB 2; Length 584;
Best Local Similarity 30.4%; Pred. No. 1.7e-33;
Matches 158; Conservative 74; Mismatches 200; Indels 88; Gaps 18;

```

DB 135 VDVA-----EIGRAFAVMSBDTDLTFRKTSQPVHIEIKFVSESHGDGAFD----- 183  
 QY 199 MAGQAWRTPLP-RAGEAHFQDDEWMSRRGRGLFVLVAHETGHTGLTHSPAPAL 257  
 DB 184 --QGGLTAAFFVFGGDHFDALMTGSPRGNTLFFVAAHEGSHGSHSQSSAL 241  
 QY 258 MAPYKRLGRDALLSDWDVLAVQSLYKPLGGSVAVOLPGKLFTEFTWDSYSPQGRAP 317  
 DB 242 MAPFYGFEPFVFLDDEDDKAAIQSLYGRKTN-----QLRPTNVYATIQRPYSP----- 290  
 QY 318 TGGPK-----YC-HSSPDAITVDQOQLYFKGSHFMEVADGVSE--PPQLQERNVG 368  
 DB 291 ---PKVPLDSDICQKSKVTLTNSAQGETYAKGKXYKLTLD--SVEEGYQQLISKWPG 346  
 QY 369 LPENTIAAAVSLNDGDFYFFKGCRCWRFRCPKPVGVLPGOLCRAG--GLPRHPDALPRP 426  
 DB 347 LRGNTD-AAFTYKNGKTYFFKCTQYWRQGRMDGVYKREISEGTTGIPDHLDAWVG 405  
 QY 427 LRLLLFKGARYYTLARGGL-OVEPYPRSLQDWGIGPEVSGALPRDGSIIFFRDY 485  
 DB 406 NCKIYFFKSGKFWRPDPKRPVKASYPKPISNMWGVNNILDALKYTNGYTFYFKGDY 465  
 QY 466 WRLOAKLQAT-----SGRMATELPWGCWTHANSGS 517  
 DB 466 YRFHDARPAVDSATPPPRPTAHM-----WRCCKNTPSST 500

## RESULT 13

Q9W122 PRELIMINARY: PRT: 613 AA.  
 AC Q9W122  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE CG4859-PA.  
 GN Name=Mmp1; ORFNames=CG4859;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OK NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer V.G., Chame M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
 RA Baller R.M., Bau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Foele C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibbegan C.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liao X., Lele Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mantei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Morkulov G., Mitsuhashi N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Chape M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskaas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCh0079-RESEARCh0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.,  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 a genomic perspective.";  
 RL Genome Biol. 3:RESEARCh0084-RESEARCh0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Beltencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review.";  
 RL Genome Biol. 3:RESEARCh0083-RESEARCh0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG Flybase;  
 RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
 RP [6]  
 RP SEQUENCE FROM N.A.  
 RG Flybase;  
 RL Submitted (MAR-2004) to the EMBL/Genbank/DBJ databases.  
 DR EMBL, AE003464; AAF47255.2; -.  
 DR HSSP, P03956; ICGL.  
 DR MEROPS, M10.031; -.  
 DR Flybase, FBgn0035049; Mmp1.  
 DR GO, GO:0004222; F:metalloendopeptidase activity; TAS.  
 DR GO, GO:0048102; P:autophagic cell death; IMP.  
 DR GO, GO:0035001; P:dorsal trunk growth; IMP.  
 DR GO, GO:0002168; P:larval development (sensu Insecta); IMP.  
 DR GO, GO:0035071; P:salivary gland cell death; IMP.  
 DR InterPro, IPR000585; Hemopexin.  
 DR InterPro, IPR006026; Peptidase\_M.  
 DR InterPro, IPR001818; Pept\_M10A\_M12B.  
 DR InterPro, IPR006025; Pept\_M\_Zn\_B8.  
 DR InterPro, IPR009070; PGSD\_like.  
 DR Pfam, PF00045; Hemopexin; 4.  
 DR Pfam, PF03933; peptidase\_M10\_N; 1.  
 DR PRINTS, PR00138; MARRIXIN.  
 DR SMART, SM00120; HX; 4.  
 DR SMART, SM00235; ZINC; 1.  
 DR PROSITE, PS00024; HEMOPEXIN; 1.  
 DR PROSITE, PS00142; ZINC\_PROTEASE; UNKNOWN 1.  
 DR SEQUENCE 613 AA; 67805 MW; EADE777AF43948B0 CRC64;



DR HSSP; P08254; I88Y.  
DR GO; GO:0005578; C:extracellular matrix (sensu Metacoa); IEA.  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2005, 17:27:41 ; Search time 43 Seconds  
(without alignments)  
902.733 Million cell updates/sec

Title: US-10-791-980-6

Perfect score: 2834  
Sequence: 1 MVARVGLLRALQLLMGHL.....WATELPMWGMWANSALF 520

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA:  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfilltest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID                | Description        |
|------------|--------|-------------|--------|-------------------|--------------------|
| 1          | 2834   | 100.0       | 520    | US-09-862-631-6   | Sequence 6, Appli  |
| 2          | 2472   | 87.2        | 520    | US-09-950-510-2   | Sequence 2, Appli  |
| 3          | 2472   | 87.2        | 520    | US-09-950-510-21  | Sequence 21, Appli |
| 4          | 2472   | 87.2        | 520    | US-09-950-510-24  | Sequence 24, Appli |
| 5          | 2444   | 86.2        | 520    | US-09-391-104-10  | Sequence 10, Appli |
| 6          | 2324.5 | 82.0        | 445    | US-09-950-510-12  | Sequence 12, Appli |
| 7          | 2207   | 77.9        | 469    | US-09-950-510-23  | Sequence 23, Appli |
| 8          | 1731   | 61.1        | 393    | US-09-950-510-22  | Sequence 22, Appli |
| 9          | 546    | 19.3        | 508    | US-09-171-545-1   | Sequence 1, Appli  |
| 10         | 543    | 19.2        | 508    | US-09-391-104-18  | Sequence 18, Appli |
| 11         | 542    | 19.1        | 517    | US-09-171-545-2   | Sequence 2, Appli  |
| 12         | 535.5  | 18.9        | 607    | US-09-000-041A-2  | Sequence 2, Appli  |
| 13         | 535.5  | 18.9        | 607    | US-09-734-002-2   | Sequence 2, Appli  |
| 14         | 534.5  | 18.9        | 607    | US-09-211-704A-10 | Sequence 10, Appli |
| 15         | 534    | 18.8        | 669    | US-08-704-711A-3  | Sequence 3, Appli  |
| 16         | 534    | 18.8        | 669    | US-09-521-220-3   | Sequence 29, Appli |
| 17         | 534    | 18.8        | 669    | US-09-391-104-29  | Sequence 2, Appli  |
| 18         | 514.5  | 18.2        | 582    | US-08-704-711A-2  | Sequence 1, Appli  |
| 19         | 514.5  | 18.2        | 582    | US-08-448-489-1   | Sequence 9, Appli  |
| 20         | 514.5  | 18.2        | 582    | US-09-211-704A-9  | Sequence 2, Appli  |
| 21         | 514.5  | 18.2        | 582    | US-09-521-220-2   | Sequence 2, Appli  |
| 22         | 514.5  | 18.2        | 582    | US-09-391-104-28  | Sequence 28, Appli |
| 23         | 514.5  | 18.2        | 582    | US-09-919-497-84  | Sequence 8, Appli  |
| 24         | 514.5  | 18.2        | 582    | US-09-689-730-1   | Sequence 1, Appli  |
| 25         | 513.5  | 18.1        | 532    | US-09-294-841-2   | Sequence 2, Appli  |
| 26         | 498.5  | 17.6        | 604    | US-09-391-104-30  | Sequence 30, Appli |
| 27         | 485    | 17.1        | 562    | US-09-211-704A-4  | Sequence 4, Appli  |

|    |       |      |     |   |                     |                    |
|----|-------|------|-----|---|---------------------|--------------------|
| 28 | 485   | 17.1 | 570 | 4 | US-09-949-016-8263  | Sequence 8263, Ap  |
| 29 | 485   | 17.1 | 570 | 4 | US-09-949-016-8264  | Sequence 8264, Ap  |
| 30 | 478.5 | 16.9 | 563 | 3 | US-09-211-704A-2    | Sequence 2, Appli  |
| 31 | 472   | 16.7 | 564 | 3 | US-09-211-704A-8    | Sequence 8, Appli  |
| 32 | 472   | 16.7 | 579 | 3 | US-08-704-711A-1    | Sequence 1, Appli  |
| 33 | 472   | 16.7 | 579 | 3 | US-09-521-220-1     | Sequence 1, Appli  |
| 34 | 466.5 | 16.5 | 492 | 1 | US-07-794-393-4     | Sequence 4, Appli  |
| 35 | 466.5 | 16.5 | 492 | 1 | US-08-001-711-4     | Sequence 4, Appli  |
| 36 | 455.5 | 16.1 | 488 | 1 | US-07-794-393-2     | Sequence 2, Appli  |
| 37 | 455.5 | 16.1 | 488 | 1 | US-08-001-711-2     | Sequence 2, Appli  |
| 38 | 455.5 | 16.1 | 488 | 3 | US-08-704-711A-22   | Sequence 22, Appli |
| 39 | 455.5 | 16.1 | 488 | 3 | US-09-521-220-22    | Sequence 22, Appli |
| 40 | 455.5 | 16.1 | 488 | 3 | US-09-391-104-31    | Sequence 31, Appli |
| 41 | 455.5 | 16.1 | 489 | 3 | US-08-448-489-11    | Sequence 11, Appli |
| 42 | 455.5 | 16.1 | 489 | 4 | US-09-689-730-11    | Sequence 11, Appli |
| 43 | 455.5 | 16.1 | 491 | 4 | US-09-949-016-10808 | Sequence 10808, A  |
| 44 | 449   | 15.8 | 416 | 4 | US-09-171-545-3     | Sequence 3, Appli  |
| 45 | 449   | 15.8 | 416 | 4 | US-09-171-545-4     | Sequence 4, Appli  |

ALIGNMENTS

|  |     |   |     |  |  |  |  |  |  |
|--|-----|---|-----|--|--|--|--|--|--|
| RESULT 1   |     |   |     |  |  |  |  |  |  |
| US-09-862-631-6  |     |   |     |  |  |  |  |  |  |
| Sequence 6, Application US/09862631                          |     |   |     |  |  |  |  |  |  |
| Patent No. 6734005   |     |   |     |  |  |  |  |  |  |
| GENERAL INFORMATION:   |     |   |     |  |  |  |  |  |  |
| APPLICANT: Holmgren, Erik                                    |     |   |     |  |  |  |  |  |  |
| APPLICANT: Kihlen, Mats                                      |     |   |     |  |  |  |  |  |  |
| APPLICANT: Wood, Tim   |     |   |     |  |  |  |  |  |  |
| APPLICANT: Ekblom, Jonas                                     |     |   |     |  |  |  |  |  |  |
| TITLE OF INVENTION: No. 6734005el Matrix Metalloproteinases  |     |   |     |  |  |  |  |  |  |
| FILE REFERENCE: 00014regus                                   |     |   |     |  |  |  |  |  |  |
| CURRENT APPLICATION NUMBER: US/09/862,631                    |     |   |     |  |  |  |  |  |  |
| CURRENT FILING DATE: 2000-05-22                              |     |   |     |  |  |  |  |  |  |
| PRIOR APPLICATION NUMBER: 206119                             |     |   |     |  |  |  |  |  |  |
| NUMBER OF SEQ ID NOS: 8                                      |     |   |     |  |  |  |  |  |  |
| SOFTWARE: PatentIn version 3.0                               |     |   |     |  |  |  |  |  |  |
| SEQ ID NO 6  |     |   |     |  |  |  |  |  |  |
| LENGTH: 520  |     |   |     |  |  |  |  |  |  |
| TYPE: PRT  |     |   |     |  |  |  |  |  |  |
| ORGANISM: Homo sapiens                                       |     |   |     |  |  |  |  |  |  |
| US-09-862-631-6  |     |   |     |  |  |  |  |  |  |
| Query Match  |     |   |     |  |  |  |  |  |  |
| Best Local Similarity 100.0%; Score 2834; DB 4; Length 520;  |     |   |     |  |  |  |  |  |  |
| Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |     |   |     |  |  |  |  |  |  |
| QY   | 1   | MVARVGLLRALQLLMGHLDAQPAERGGQELRKEAFAFLKRYGYLNEQVPAPTSTRFS   | 60  |  |  |  |  |  |  |
| DB   | 1   | MVARVGLLRALQLLMGHLDAQPAERGGQELRKEAFAFLKRYGYLNEQVPAPTSTRFS   | 60  |  |  |  |  |  |  |
| QY   | 61  | DAIRAFQWVSQLPVSGVLDRAITLRQTRPRCGVTDTNSYAAMAEIRISDLFAHRTKMRK | 120 |  |  |  |  |  |  |
| DB   | 61  | DAIRAFQWVSQLPVSGVLDRAITLRQTRPRCGVTDTNSYAAMAEIRISDLFAHRTKMRK | 120 |  |  |  |  |  |  |
| QY   | 121 | KRFPAQGNKWKYKQHLSTYLVNMPETLRBQRGAPAPAPSSCGATQRRSSGRRPOQAPLT | 180 |  |  |  |  |  |  |
| DB   | 121 | KRFPAQGNKWKYKQHLSTYLVNMPETLRBQRGAPAPAPSSCGATQRRSSGRRPOQAPLT | 180 |  |  |  |  |  |  |
| QY   | 181 | SGSPSSKGTITTMGMAMPMAOGAPRPTFFLPARGAHPDODERWLSRRRGNTLFVVLAH  | 240 |  |  |  |  |  |  |
| DB   | 181 | SGSPSSKGTITTMGMAMPMAOGAPRPTFFLPARGAHPDODERWLSRRRGNTLFVVLAH  | 240 |  |  |  |  |  |  |
| QY   | 241 | EIGHTLGLTSPAPRALAPYKRLGRDALLSWDDVLAVQSLYKRLGGSVAVALPGKLF    | 300 |  |  |  |  |  |  |
| DB   | 241 | EIGHTLGLTSPAPRALAPYKRLGRDALLSWDDVLAVQSLYKRLGGSVAVALPGKLF    | 300 |  |  |  |  |  |  |
| QY   | 301 | TDFFWDSYSGQRRPETQGRKYCHSSPDATTVDRQOOLYIFKSHFWEVADGVSEPR     | 360 |  |  |  |  |  |  |
| DB   | 301 | TDFFWDSYSGQRRPETQGRKYCHSSPDATTVDRQOOLYIFKSHFWEVADGVSEPR     | 360 |  |  |  |  |  |  |

QY 361 PLOERWVGLPNTIEAAVSLNDGFYFFKGRGCMRFGPKPVWGLPOLCRAGGLPRHPDA 420  
DB 361 PLOERWVGLPNTIEAAVSLNDGFYFFKGRGCMRFGPKPVWGLPOLCRAGGLPRHPDA 420  
QY 421 ALFFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGIPBESVSGALPRPDGSIIF 480  
DB 421 ALFFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGIPBESVSGALPRPDGSIIF 480  
QY 481 RDDRWRDLQAKLQATTSGRWATELPMWGMCHANSGLALF 520  
DB 481 RDDRWRDLQAKLQATTSGRWATELPMWGMCHANSGLALF 520

RESULT 2  
US-09-950-510-2

/ Sequence 2, Application US/09950510  
/ Patent No. 6740514  
/ GENERAL INFORMATION:  
/ APPLICANT: Curtis, Rory  
/ TITLE OF INVENTION: 46798, A No. 6740514e1 Human Matrix Metalloproteinase And Uses Th  
/ FILE REFERENCE: 10147-45U1  
/ CURRENT APPLICATION NUMBER: US/09/950,510  
/ PRIOR FILING DATE: 2001-09-10  
/ PRIOR APPLICATION NUMBER: 60/251,156  
/ NUMBER OF SEQ ID NOS: 24  
/ SOFTWARE: PatentIn version 3.0  
/ SEQ ID NO 2  
/ LENGTH: 520  
/ TYPE: PRF  
/ ORGANISM: Homo sapiens  
US-09-950-510-2

Query Match 87.2%; Score 2472; DB 4; Length 520;  
Best Local Similarity 88.9%; Pred. No. 8.9e-238;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRLALQLLMGHLDQPAERGOELRKEAEAFLEKGYLNEQVPAKPTSTRFS 60  
DB 1 MVARVGLLRLALQLLMGHLDQPAERGOELRKEAEAFLEKGYLNEQVPAKPTSTRFS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTTRPCGVTDTNSYAAMERISDLFAHRTKMRK 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTTRPCGVTDTNSYAAMERISDLFAHRTKMRK 120  
QY 121 KRPAKQGNKMYKQHLSTRLVNMPHLSRQFGAPCAPSSCGATS--QRWS--SGRPQ 175  
DB 121 KRPAKQGNKMYKQHLSTRLVNMPHLSRQFGAPCAPSSCGATS--QRWS--SGRPQ 175  
QY 176 QAPLTSGPS--KCTTGMWAMPMAOGAEWRTPFLPRRGEAHFDODERWSLSRR 229  
DB 176 QAPLTSGPS--KCTTGMWAMPMAOGAEWRTPFLPRRGEAHFDODERWSLSRR 229  
QY 230 RGRNLFVYLAHEIGTIGLTHSPAPRALMAPYKRLGSDALLSMDVYLAOSLYGKPLGG 289  
DB 230 RGRNLFVYLAHEIGTIGLTHSPAPRALMAPYKRLGSDALLSMDVYLAOSLYGKPLGG 289  
QY 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDATYVDRQOOLYIFKGSHPWE 349  
DB 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDATYVDRQOOLYIFKGSHPWE 349  
QY 350 VAAAGNVSEPRPLOERWVGLPNTIEAAVSLNDGFYFFKGRGCMRFGPKPVWGLPOLC 409  
DB 350 VAAAGNVSEPRPLOERWVGLPNTIEAAVSLNDGFYFFKGRGCMRFGPKPVWGLPOLC 409  
QY 410 RAGGLPRHPDALFFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGIPBESVGA 469  
DB 410 RAGGLPRHPDALFFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGIPBESVGA 469  
QY 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGMCHANSGLALF 520  
DB 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGMCHANSGLALF 520

RESULT 3  
US-09-950-510-21

/ Sequence 21, Application US/09950510  
/ Patent No. 6740514  
/ GENERAL INFORMATION:  
/ APPLICANT: Curtis, Rory  
/ TITLE OF INVENTION: 46798, A No. 6740514e1 Human Matrix Metalloproteinase And Uses Th  
/ FILE REFERENCE: 10147-45U1  
/ CURRENT APPLICATION NUMBER: US/09/950,510  
/ PRIOR FILING DATE: 2001-09-10  
/ PRIOR APPLICATION NUMBER: 60/251,156  
/ NUMBER OF SEQ ID NOS: 24  
/ SOFTWARE: PatentIn version 3.0  
/ SEQ ID NO 21  
/ LENGTH: 520  
/ TYPE: PRF  
/ ORGANISM: Homo sapiens  
US-09-950-510-21

Query Match 87.2%; Score 2472; DB 4; Length 520;  
Best Local Similarity 88.9%; Pred. No. 8.9e-238;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRLALQLLMGHLDQPAERGOELRKEAEAFLEKGYLNEQVPAKPTSTRFS 60  
DB 1 MVARVGLLRLALQLLMGHLDQPAERGOELRKEAEAFLEKGYLNEQVPAKPTSTRFS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTTRPCGVTDTNSYAAMERISDLFAHRTKMRK 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAATLRQMTTRPCGVTDTNSYAAMERISDLFAHRTKMRK 120  
QY 121 KRPAKQGNKMYKQHLSTRLVNMPHLSRQFGAPCAPSSCGATS--QRWS--SGRPQ 175  
DB 121 KRPAKQGNKMYKQHLSTRLVNMPHLSRQFGAPCAPSSCGATS--QRWS--SGRPQ 175  
QY 176 QAPLTSGPS--KCTTGMWAMPMAOGAEWRTPFLPRRGEAHFDODERWSLSRR 229  
DB 176 QAPLTSGPS--KCTTGMWAMPMAOGAEWRTPFLPRRGEAHFDODERWSLSRR 229  
QY 230 RGRNLFVYLAHEIGTIGLTHSPAPRALMAPYKRLGSDALLSMDVYLAOSLYGKPLGG 289  
DB 230 RGRNLFVYLAHEIGTIGLTHSPAPRALMAPYKRLGSDALLSMDVYLAOSLYGKPLGG 289  
QY 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDATYVDRQOOLYIFKGSHPWE 349  
DB 290 SVAVOLPGKLTDFETWDSYSPQGRRPETQPKYCHSSFDATYVDRQOOLYIFKGSHPWE 349  
QY 350 VAAAGNVSEPRPLOERWVGLPNTIEAAVSLNDGFYFFKGRGCMRFGPKPVWGLPOLC 409  
DB 350 VAAAGNVSEPRPLOERWVGLPNTIEAAVSLNDGFYFFKGRGCMRFGPKPVWGLPOLC 409  
QY 410 RAGGLPRHPDALFFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGIPBESVGA 469  
DB 410 RAGGLPRHPDALFFPPLRLILFKGARYVYLARGGLQVEPYPRSLQDWGIPBESVGA 469  
QY 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGMCHANSGLALF 520  
DB 470 LPRPDGSIIFPRDDRWRDLQAKLQATTSGRWATELPMWGMCHANSGLALF 520

## RESULT 4

US-09-950-510-24  
/ Sequence 24, Application US/09950510  
/ Patent No. 6740514  
/ GENERAL INFORMATION:  
/ APPLICANT: Curtis, Rory  
/ TITLE OF INVENTION: 46798, A No. 6740514e1 Human Matrix Metalloproteinase And Uses Th  
/ FILE REFERENCE: 10147-45U1  
/ CURRENT APPLICATION NUMBER: US/09/950,510  
/ PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/251,156  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 24  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-950-510-24

Query Match 87.2%; Score 2472; DB 4; Length 520;  
Best Local Similarity 88.9%; Pred. No. 8.9e-238;  
Matches 472; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQQLLMGHLDAQPAERGGQELRKEAEFLKGYGLNEQVPKAPTSTRS 60  
DB 1 MVARVGLLRALQQLLMGHLDAQPAERGGQELRKEAEFLKGYGLNEQVPKAPTSTRS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAQLRQMTTRPCGVTDNTSYAAMAEKISDLFAHRTKMRK 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAQLRQMTTRPCGVTDNTSYAAMAEKISDLFAHRTKMRK 120  
QY 121 KRFPAQGNKWKYKQHLSTYRLVNMPEHLRSQFGAPCAPSSCGATS---QRMS--SGRPOP 175  
DB 121 KRFPAQGNKWKYKQHLSTYRLVNMPEHLRSQFGAPCAPSSCGATS---QRMS--SGRPOP 175  
QY 176 QAPLTSGSPSS-----KGTITMGWAMPMAQGAEMRTPEFLPRGGAHFDODERMSLSRR 229  
DB 176 QAPLTSGSPSS-----KGTITMGWAMPMAQGAEMRTPEFLPRGGAHFDODERMSLSRR 229  
QY 230 RGRNLFVVLAAHEIGTGLTSPAPRALMAPYKRLGRDALISWDVLAQSLYKGPILG 289  
DB 230 RGRNLFVVLAAHEIGTGLTSPAPRALMAPYKRLGRDALISWDVLAQSLYKGPILG 289  
QY 290 SVAVOLPGKLFDFEFTWDSYSPQGRPEPTQGGKCHSSFDALTVDROOOLYFKGSHFWE 349  
DB 290 SVAVOLPGKLFDFEFTWDSYSPQGRPEPTQGGKCHSSFDALTVDROOOLYFKGSHFWE 349  
QY 350 VAADGNVSEPRPQGRWGLPNIIEAAVSLNDGDFYFFKGGRCMRFRGPKVWGILPOLC 409  
DB 350 VAADGNVSEPRPQGRWGLPNIIEAAVSLNDGDFYFFKGGRCMRFRGPKVWGILPOLC 409  
QY 410 RAGGLPRHPDALLFPPLRLILFKGARYVVLARGLQVEPYPPSLDQWGIPEVSGA 469  
DB 410 RAGGLPRHPDALLFPPLRLILFKGARYVVLARGLQVEPYPPSLDQWGIPEVSGA 469  
QY 470 LRPDGSIIFFRDRIYRLDQAKLQATTSGRWATELPMWGMHANSGLALF 520  
DB 470 LRPDGSIIFFRDRIYRLDQAKLQATTSGRWATELPMWGMHANSGLALF 520

## RESULT 5

US-09-391-104-10  
Sequence 10, Application US/09391104  
Patent No. 639371  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Falduto, Michael T.  
APPLICANT: Magnuson, Scott R.  
APPLICANT: Morgan, Douglas W.  
TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEINASE GENE,  
TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS  
TITLE OF INVENTION: OF USING SAME  
FILE REFERENCE: 6073.US.21  
CURRENT APPLICATION NUMBER: US/09/391,104  
CURRENT FILING DATE: 1999-09-07  
PRIOR APPLICATION NUMBER: US 08/814,394  
PRIOR FILING DATE: 1997-03-11  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 520  
TYPE: PRT

ORGANISM: Homo sapiens  
US-09-391-104-10

Query Match 86.2%; Score 2444; DB 3; Length 520;  
Best Local Similarity 87.9%; Pred. No. 5.6e-235;  
Matches 467; Conservative 7; Mismatches 35; Indels 22; Gaps 5;

QY 1 MVARVGLLRALQQLLMGHLDAQPAERGGQELRKEAEFLKGYGLNEQVPKAPTSTRS 60  
DB 1 MVARVGLLRALQQLLMGHLDAQPAERGGQELRKEAEFLKGYGLNEQVPKAPTSTRS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAQLRQMTTRPCGVTDNTSYAAMAEKISDLFAHRTKMRK 120  
DB 61 DAIRAFQWVSQLPVSGVLDRAQLRQMTTRPCGVTDNTSYAAMAEKISDLFAHRTKMRK 120  
QY 121 KRFPAQGNKWKYKQHLSTYRLVNMPEHLRSQFGAPCAPSSCGATS---QRMS--SGRPOP 175  
DB 121 KRFPAQGNKWKYKQHLSTYRLVNMPEHLRSQFGAPCAPSSCGATS---QRMS--SGRPOP 175  
QY 176 QAPLTSGSPSS-----KGTITMGWAMPMAQGAEMRTPEFLPRGGAHFDODERMSLSRR 229  
DB 176 QAPLTSGSPSS-----KGTITMGWAMPMAQGAEMRTPEFLPRGGAHFDODERMSLSRR 229  
QY 230 RGRNLFVVLAAHEIGTGLTSPAPRALMAPYKRLGRDALISWDVLAQSLYKGPILG 289  
DB 230 RGRNLFVVLAAHEIGTGLTSPAPRALMAPYKRLGRDALISWDVLAQSLYKGPILG 289  
QY 290 SVAVOLPGKLFDFEFTWDSYSPQGRPEPTQGGKCHSSFDALTVDROOOLYFKGSHFWE 349  
DB 290 SVAVOLPGKLFDFEFTWDSYSPQGRPEPTQGGKCHSSFDALTVDROOOLYFKGSHFWE 349  
QY 350 VAADGNVSEPRPQGRWGLPNIIEAAVSLNDGDFYFFKGGRCMRFRGPKVWGILPOLC 409  
DB 350 VAADGNVSEPRPQGRWGLPNIIEAAVSLNDGDFYFFKGGRCMRFRGPKVWGILPOLC 409  
QY 410 RAGGLPRHPDALLFPPLRLILFKGARYVVLARGLQVEPYPPSLDQWGIPEVSGA 469  
DB 410 RAGGLPRHPDALLFPPLRLILFKGARYVVLARGLQVEPYPPSLDQWGIPEVSGA 469  
QY 470 LRPDGSIIFFRDRIYRLDQAKLQATTSGRWATELPMWGMHANSGLALF 520  
DB 470 LRPDGSIIFFRDRIYRLDQAKLQATTSGRWATELPMWGMHANSGLALF 520

## RESULT 6

US-09-950-510-12  
Sequence 12, Application US/09950510  
Patent No. 6740514  
GENERAL INFORMATION:  
APPLICANT: Curis, Rory  
TITLE OF INVENTION: 46798, A No. 6740514el Human Matrix Metalloproteinase And Uses Th  
FILE REFERENCE: 10147-45U1  
CURRENT APPLICATION NUMBER: US/09/950,510  
CURRENT FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/251,156  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 445  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-950-510-12

Query Match 82.0%; Score 2324.5; DB 4; Length 445;  
Best Local Similarity 84.8%; Pred. No. 3.7e-223;  
Matches 441; Conservative 1; Mismatches 3; Indels 75; Gaps 2;

QY 1 MVARVGLLRALQQLLMGHLDAQPAERGGQELRKEAEFLKGYGLNEQVPKAPTSTRS 60  
DB 1 MVARVGLLRALQQLLMGHLDAQPAERGGQELRKEAEFLKGYGLNEQVPKAPTSTRS 60  
QY 61 DAIRAFQWVSQLPVSGVLDRAQLRQMTTRPCGVTDNTSYAAMAEKISDLFAHRTKMRK 120

```

Db      61 DAIRAFQWVSQLPVSGVLDRAITLROMTSPRCGVTDTNSYAAMERISDLFARHRTKRRK 120
Qy      121 KRPAKQGNKWKYKQHLSTYLVNMPBHLSRQFGACAPSSSCGANSQWSSGRPQAPLPT 180
Db      121 KRPAKQG-----GALAH-----132
Qy      181 SGSSSKGTTTGMAMPMLAQAAPWRTPELPRGGAHFDDERMSLSRRGRNLFVTLAH 240
Db      133 -----AFLEPRGEAHFDODERMSLSRRGRNLFVTLAH 165
Qy      241 EIGHTLGLTHSPARALMAPYKRLGRDALISWDVLAQVSLYKPIFGSSVAVOLPGKLF 300
Db      166 EIGHTLGLTHSPARALMAPYKRLGRDALISWDVLAQVSLYKPIFGSSVAVOLPGKLF 225
Qy      301 TDFETWDSYSPQGRRPETQGPKYCHSSFDAITVDROOQLYIFKGSHPWEVAADGANSEPP 360
Db      226 TDFETWDSYSPQGRRPETQGPKYCHSSFDAITVDROOQLYIFKGSHPWEVAADGANSEPP 285
Qy      361 PLOERWVGLPNNIEAALVSLNDGDFYFKGRCWRFGRPKVWGLPOLCRAGGLPRHPDA 420
Db      286 PLOERWVGLPNNIEAALVSLNDGDFYFKGRCWRFGRPKVWGLPOLCRAGGLPRHPDA 345
Qy      421 ALFEPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGIDEVSGALPRPDGSIIFP 480
Db      346 ALFEPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWGIDEVSGALPRPDGSIIFP 405
Qy      481 RDRYWRDLQAKLQATTSGRWATELPMWGCMAHNSGALF 520
Db      406 RDRYWRDLQAKLQATTSGRWATELPMWGCMAHNSGALF 445

```

## RESULT 7

```

US-09-950-510-23
; Sequence 23, Application US/09950510
; Patent No. 6740514
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy
; TITLE OF INVENTION: 46798, A No. 6740514el Human Matrix Metalloproteinase And Uses Th
; FILE REFERENCE: 10147-45U1
; CURRENT APPLICATION NUMBER: US/09/950,510
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/251,156
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-510-23

```

```

Query Match      77.9%; Score 2207; DB 4; Length 469;
Best Local Similarity 87.5%; Pred. No. 2,2e-211;
Matches 420; Conservative 4; Mismatches 34; Indels 22; Gaps 5;

Qy      52 KAPSTRFSDAIRAFQWVSQLPVSGVLDRAITLROMTSPRCGVTDTNSYAAMERISDLFA 111
Db      1 KAPSTRFSDAIRAFQWVSQLPVSGVLDRAITLROMTSPRCGVTDTNSYAAMERISDLFA 60
Qy      112 RHRTRKRRKRPAAKQGNKWKYKQHLSTYLVNMPBHLSRQFGACAPSSSCGATS---QRW 168
Db      61 RHRTRKRRKRPAAKQGNKWKYKQHLSTYLVNMPBHLSRQFGACAPSSSCGATS---QRW 111
Qy      169 S--SGRPPQAPLVTSGSPSS-----KGTTMGMAMPMLAQAAPWRTPELPRGGAHFDD 220
Db      112 SNVSLALFEWEAPAT--GPADIRLITFFQGDHNDGLGNAPDGPGLAHAFLEPRRGEAHFDQ 169
Qy      221 DERWLSLSRRGRNLFVTLAHEIGHTLGLTHSPARALMAPYKRLGRDALISWDVLAQV 280
Db      170 DERWLSLSRRGRNLFVTLAHEIGHTLGLTHSPARALMAPYKRLGRDALISWDVLAQV 229
Qy      281 SLYKPIFGSSVAVOLPGKLFDTFETWDSYSPQGRRPETQGPKYCHSSFDAITVDROOQLY 340

```

```

Db      230 SLYKPIFGSSVAVOLPGKLFDTFETWDSYSPQGRRPETQGPKYCHSSFDAITVDROOQLY 289
Qy      341 IFKGSHPWEVAADGANVSEBRPLQERWVGLPNNIEAALVSLNDGDFYFKGRCWRFGRPK 400
Db      290 IFKGSHPWEVAADGANVSEBRPLQERWVGLPNNIEAALVSLNDGDFYFKGRCWRFGRPK 349
Qy      401 PYWGLPOLCRAGGLPRHPDAALFEPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWG 460
Db      350 PYWGLPOLCRAGGLPRHPDAALFEPPLRLILFKGARYYVLARGGLQVEPYPRSLQDWG 409
Qy      461 GIPREVSGALPRPDGSIIFRDDRWRDLQAKLQATTSGRWATELPMWGCMAHNSGALF 520
Db      410 GIPREVSGALPRPDGSIIFRDDRWRDLQAKLQATTSGRWATELPMWGCMAHNSGALF 469

```

## RESULT 8

```

US-09-950-510-22
; Sequence 22, Application US/09950510
; Patent No. 6740514
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy
; TITLE OF INVENTION: 46798, A No. 6740514el Human Matrix Metalloproteinase And Uses Th
; FILE REFERENCE: 10147-45U1
; CURRENT APPLICATION NUMBER: US/09/950,510
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/251,156
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-950-510-22

```

```

Query Match      61.1%; Score 1731; DB 4; Length 393;
Best Local Similarity 85.2%; Pred. No. 5.3e-164;
Matches 341; Conservative 4; Mismatches 33; Indels 22; Gaps 5;

Qy      1 MVARVGLLRLALQLLLMGHLDAPAEKGGQELKKAFAFLKKGYNLEQVPAKPTSTRFS 60
Db      1 MVARVGLLRLALQLLLMGHLDAPAEKGGQELKKAFAFLKKGYNLEQVPAKPTSTRFS 60
Qy      61 DAIRAFQWVSQLPVSGVLDRAITLROMTSPRCGVTDTNSYAAMERISDLFARHRTKRRK 120
Db      61 DAIRAFQWVSQLPVSGVLDRAITLROMTSPRCGVTDTNSYAAMERISDLFARHRTKRRK 120
Qy      121 KRPAKQGNKWKYKQHLSTYLVNMPBHLSRQFGACAPSSSCGATS---QRW--SGRPPQ 175
Db      121 KRPAKQGNKWKYKQHLSTYLVNMPBHLSRQFGACAPSSSCGATS---QRW--SGRPPQ 175
Qy      176 QAPLTSGPS-----KGTTMGMAMPMLAQAAPWRTPELPRGGAHFDDERMSLSRR 229
Db      172 EAPAT--GPADIRLITFFQGDHNDGLGNAPDGPGLAHAFLEPRRGEAHFDQ 229
Qy      230 RGNLFLVTLAHEIGHTLGLTHSPARALMAPYKRLGRDALISWDVLAQVSLYKPIFG 289
Db      230 RGNLFLVTLAHEIGHTLGLTHSPARALMAPYKRLGRDALISWDVLAQVSLYKPIFG 289
Qy      290 SVAVOLPGKLFDTFETWDSYSPQGRRPETQGPKYCHSSFDAITVDROOQLYIFKGSHPWE 349
Db      290 SVAVOLPGKLFDTFETWDSYSPQGRRPETQGPKYCHSSFDAITVDROOQLYIFKGSHPWE 349
Qy      350 VAADGANVSEBRPLQERWVGLPNNIEAALVSLNDGDFYFKP 389
Db      350 VAADGANVSEBRPLQERWVGLPNNIEAALVSLNDGDFYFKP 389

```

## RESULT 9

```

US-09-171-545-1
; Sequence 1, Application US/09171545
; Patent No. 656116

```



```

Oy 24 PAREGGQGLERKAAE-FLKXGYLNEQVXPAPSTISRED---AIRAQWMSQLPVSGLD 79
Db 23 PEE-----KEAVDYLLQYGLQKPLEGA-DDFRLEDITEALRTQEASELFPVSGMD 74
Oy 80 RATTLOMTRPRCGVTDTNSYAAAMERISDLFAHRTKMRKRKRAKQGNKYKOHLSYRL 139
Db 75 DATRAAMQPRCGLEDPFN-----QKTLKULLLGH-WRKKHLTFRI 114
Oy 140 VVMPHILASRQGAFCAPSSCGATSQWSSGRRPOPAPLTSSGSPSSKGTTHMGWA-MPL 198
Db 115 LVNPFSTL-----SPSRVRAALHQAFKYSN-----VAPLTFREVKA-----GMADIRL 157
Oy 199 MAQAPMPPTF-----LPRGGAHPDODERWSLSRRRGHNLFWVLAAHET 242
Db 158 SFHGR--QSPFCSNSFDPRGKVLAAHADVPELGSVHFNDDEFWETGYQVNLHIAHEV 215
Oy 243 GHTGLTSPAPRALMAFYRRLGRDALLSMDVLAVQSLYKPELGSVAVQLPGKLTFD 302
Db 216 GHALGLGHSRYQLMALMAVYAGYQPYRFLHPDDVAIGLQALYK-----258
Oy 303 FETMDSYSPQRRRETQG-----PRYCHSSFDALTVDROQLYIF 342
Db 259 -----RREPEDEEBEEMHTVSTVTTKPSMPMPDCESEVDAMMLGPRGKTYAF 307
Oy 343 KGSHFWEVAAADGNVSEPRPL---QERWVGLPENTEAALVNLNODPYFFFGSGCWRFRGP 399
Db 308 KGDVWVYVTDG---RGLPRVSALWEGLPNDALVAVSPRTQRTFFFGNRYWRYVDF 363
Oy 400 KPVWGLPOLCRAGLPRHPDALFFPPLRRLILFKGARVY---VLARGGLQVEPYPRSL 456
Db 364 KLSGPFPM--KLNVHEPNLDALYWPVNOKVFLFKSGGVQWQMBELRTTDS---RFPKRI 418
Oy 457 OD-WGGIPEVVSGLPRPDGSIIFRDDRWRRLDQALQATTSGRNATELPMWGC 510
Db 419 KELTGVDPDQSAAMSWODQVYFFKGEKEYWRNLQOLRAVKGYPRWTTH--WMHC 471

RESULT 12
US-09-000-041A-2
Sequence 2, Application US/09000041A
Patent No. 6191255
GENERAL INFORMATION:
APPLICANT: Motoharu SEIKI et al.
TITLE OF INVENTION: NOVEL PROTEIN AND MONOCLONAL ANTIBODY SPECIFIC THERETO
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/000,041A
FILING DATE: January 13, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/01956
FILING DATE: July 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:

```

```

1  ? INFORMATION FOR SEQ ID NO: 2 :
2  ? SEQUENCE CHARACTERISTICS:
3  ?     LENGTH: 607
4  ?     TYPE: Amino acid
5  ?     STRANDEDNESS: Single
6  ?     TOPOLOGY: Linear
7  ?     MOLECULE TYPE: Protein
8  ?     ORIGINAL SOURCE:
9  ?         ORGANISM: Human
10 ?     US-09-000-041A-2
11
12 Query Match      18.9%, Score 535.5, DB 3, Length 607;
13 Best Local Similarity 28.7%, Pred. No. 2,2e-44;
14 Matches 158; Conservative 84; Mismatches 216; Indels 93; Gaps 22;
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
100
```



```
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,002
FILING DATE: 12-Dec-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/01956
FILING DATE: July 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 607
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Protein
ORIGINAL SOURCE:
ORGANISM: Human
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-734-002-2
```

```
Query Match 18.9%; Score 535.5; DB 4; Length 607;
Best Local Similarity 28.7%; Pred. No. 2,2e-44;
Matches 158; Conservative 84; Mismatches 216; Indels 93; Gaps 22;
```

```
12 LQLLIWHLDQAPRGGQELKEAFLAEKGYLINEOVPAKPTSTRS-----DA 62
23 LQTLIM-----ILCATVCGTEQYFNVEWVLOKGYL-----PPTDPRMSVLNSAETMOSA 72
63 IRAPQWSQLPVSGVLDRAATLRQMTBPRCGVTDNYSYAMAMERISDLFARHRTKMR-RKK 121
73 LAAMQOFGYGINNTGVNDRNTIDWKKPRCGVDPDQ-----TRGSSKHIRK 118
122 RPAKQGNKYKQHLISYRLVNW-----PEHLRS--RQFAPCAPSSCGATSQRWSSGR 173
119 RYALTGQKQKHGHITYSIKNTVPRKGDPETRKAIKRAFDV-----WQNTVP 164
174 -----QPAPLTSSGSSSKGTTM-----GWAMPLMAQAPWRTPLPR--RGAHPDQ 220
165 LTFEEVPSYLENGKRDVDITIIIFASGFHGDSSPFDGSGFLAHAYFFPGPGIGDTHFDS 224
221 DERNGLS--RRGRNLFVVLAAHEIGHTLGLTSPAPRALMAYRYKRLGSDAL-LSMDPVL 277
225 DEPWTLGNPNHNDGNLFLVAVHELGHALGLSHSNPTALMAFPYQYMETDNFKLPRNDLQ 284
278 AVQSLYG-----KPLGSAVAQVLPGLFTDFETWDSYSPQGRRPETGPKY-----CH 325
285 GIQKIYGPDKIPPRPRPLPYVPRHSIRPADPRKDKRPPRPRTGRSRYSGAKENICD 344
336 SSPDAITVDROQOLYIFKGSHPWEVAADGNVSEPRPLQER--WVGLPRNIEAAYSLNDG 383
345 GNFTLATILR-REMVFQKQWFMWRV-RNNRVMDGVQMQITYFWRGLPSPSIDAVYER-SDG 401
384 DYPFRFGSCRRFRGKPYWGLPQ--LCAAGLPRRP-DALFPPRLRLILFKARAYV 440
402 NFVFFGKRYWFKDITLQPGYPHDILITLGSGLPRPGIDSAIMWEDVGTFFGKGRYR 461
441 LARGGLQVEPYPRSLQDWGCIPEEVSQA-LPRPDGSIIFRDDRWRDLQAKLQATTSG 499
462 YSEEMKTMPPGYPKPTTWKGIPEBPGQAFVHKENGFTYFYKCKEYWKNNQILKAYEPH 521
500 RNATELPMWGC 510
522 PRSILKDFMGC 532
DB
```

```
RESULT 14
US-09-211-704A-10
Sequence 10, Application US/09211704A
```

```
Patent No. 6271014
GENERAL INFORMATION:
APPLICANT: de Saint-Vin, Blandine Marie
APPLICANT: Fossiez, Francois
APPLICANT: Caux, Christophe
APPLICANT: Lebecque, Serge J.E.
TITLE OF INVENTION: Mammalian Proteinases; Related Reagents
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESS: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/211,704A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/005,263
FILING DATE: 09-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0781K
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-9196
TELEFAX: (650) 496-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 607 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-211-704A-10
```

```
Query Match 18.9%; Score 534.5; DB 3; Length 607;
Best Local Similarity 28.7%; Pred. No. 2,8e-44;
Matches 158; Conservative 84; Mismatches 216; Indels 93; Gaps 22;
```

```
12 LQLLIWHLDQAPRGGQELKEAFLAEKGYLINEOVPAKPTSTRS-----DA 62
23 LQTLIM-----ILCATVCGTEQYFNVEWVLOKGYL-----PPTDPRMSVLNSAETMOSA 72
63 IRAPQWSQLPVSGVLDRAATLRQMTBPRCGVTDNYSYAMAMERISDLFARHRTKMR-RKK 121
73 LAAMQOFGYGINNTGVNDRNTIDWKKPRCGVDPDQ-----TRGSSKHIRK 118
122 RPAKQGNKYKQHLISYRLVNW-----PEHLRS--RQFAPCAPSSCGATSQRWSSGR 173
119 RYALTGQKQKHGHITYSIKNTVPRKGDPETRKAIKRAFDV-----WQNTVP 164
174 -----QPAPLTSSGSSSKGTTM-----GWAMPLMAQAPWRTPLPR--RGAHPDQ 220
165 LTFEEVPSYLENGKRDVDITIIIFASGFHGDSSPFDGSGFLAHAYFFPGPGIGDTHFDS 224
221 DERNGLS--RRGRNLFVVLAAHEIGHTLGLTSPAPRALMAYRYKRLGSDAL-LSMDPVL 277
225 DEPWTLGNPNHNDGNLFLVAVHELGHALGLSHSNPTALMAFPYQYMETDNFKLPRNDLQ 284
DB
```

QY 278 AVOSLYG-----KPLGSSVAVOLPGKLFDTFETWDSYSGRPRETOGPKY-----CH 325  
 Db 285 GIGKIGPPDKIIPPPRPPLPTVPSPHSIPPADRRKDRPPRPPTGRBSPGAKPNICD 344  
 QY 326 SSSDALTVDROOOLYIFKSGHFEVADGNVSEPRLOER--WVGLPNIEMAAVSLANG 383  
 Db 345 GNETLAIIR-REMFVKDQMFWRV-RNNRVMDGYMQOLTYFWRGLPSPIDAYEN-SDG 401  
 QY 384 DFFEFKGRQWRPRGPKVWGLPQ--LCRAAGLPNRP-DALFPPRLRLILFKGARYYV 440  
 Db 402 NVVFFGKNYWFVKDTTLQPGPHDLITLGGGIPPHIGIDSAIWMEDVGKTYFVKDRYWR 461  
 QY 441 IARGQLQVPPYPRSLQDWGIPPEVSGA-LPRPOSSIFPRDDRYMLDQAKLQATTSG 499  
 Db 462 YSEEMKTMDPGYPKPIITWKGILPESPOGAFVHKENGFTYFKGKEWKNQILKVEPGY 521  
 QY 500 RWATELPMWGC 510  
 Db 522 PRSLIKDPMGC 532

## RESULT 15

US-08-704-711A-3  
 ; Sequence 3, Application US/08704711A  
 ; Patent No. 6114159  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WILLY, Horst  
 ; APPLICANT: HINZMANN, Bernd  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX  
 ; NUMBER OF SEQUENCES: 22  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/704, 711A  
 ; FILING DATE: 20-NOV-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO PCT/DE95/00357  
 ; FILING DATE: 17-MAR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: DE 4438838.1  
 ; FILING DATE: 21-OCT-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: DE 4409663.1  
 ; FILING DATE: 17-MAR-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: GRANADOS, Patricia D.  
 ; REGISTRATION NUMBER: 33,683  
 ; REFERENCE/DOCKET NUMBER: 26083/124  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 672-5300  
 ; TELEFAX: (202) 672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 669 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-704-711A-3

Query Match 18.8%; Score 534; DB 3; Length 669;

Best Local Similarity 27.4%; Pred. No. 3, 6e-44;  
 Matches 167; Conservative 64; Mismatches 206; Indels 172; Gaps 21;  
 QY 9 IHALQILLMGLHDAQPAERGGGELRKEAFLPEKXYGYNLEQVPKAPTSTR-----FSDAI 63  
 Db 29 LIPLLVLLIGCIGLVV---AEDAEVHAENMLRLYGLYLPQ--PSRHMSTMRSAQLIASAL 83  
 QY 64 RAFPQVNSQLPVSGVLDRAITLQMTPRRCGVITDINSYAAMERISDLFARHRTKMRKRGF 123  
 Db 84 AEMQRFYGIPTVGLVDEETKEWKPRCGVDP--QPGRVK-----ANLRRRRRY 132  
 QY 124 AKQGNKKVKKQLSTRLVNMPHLSRQFGACAPSSCGATSSQWSSGRPOQAPLTSGS 183  
 Db 133 ALTRGKNNHLLTTSIQYTEKL----- 155  
 QY 184 PSSKQTTTMCW-----AMPLAQAQAPRTFPLPR----- 213  
 Db 156 -----GWYHSMNAVRAFRVWEQATPLVQEVPEYEDIRLRROKADIMVLFAAGFH 206  
 QY 214 -----GEAHPDDEKWSLSRR--RGRLFYVLAHEIGHITG 247  
 Db 207 GDSPPFDGTGFLAHAYFPGPLGGDTHFPADEPWTSSSTDHGNMLPLVAVHELGHALG 266  
 QY 248 LTHSPAPRALMAPYKRLGRDAL-LSWDVYLAQSLYKPLG-----GSYAVQUPGK 298  
 Db 267 LEHSSNPNALMAPPYQWKVDVNFKLPEDDLRLGIQLYGTGPDGPQPTQPLPTVTPRRPGR 326  
 QY 299 LFTDFETWDSYSPQ-----GRRETOGPKCHSPDAITV 333  
 Db 327 P-----BHRPRPPQPPPPGKPRPRPKPPVOPPRATERPDQYGPNCDDPFTVAM 379  
 QY 334 DRQOOLYIFKSGHFEVADGNVSE-PRPLQERWGLPNIEMAAVSLNDGDFPFKGR 392  
 Db 380 LR-GEMFVKGRWFMVRHNRVLDNYPMPDIGHFMRLPGDI-SAIERQDGRFVFKGR 437  
 QY 393 CWRFRGPKPVWGLPQ-LCRAG-GLP-RHPDALFPPRLRLILFKGARYYVLAAGLQVE 449  
 Db 438 YVLFREANLEPGYPOPLTSYGLIPYDRIDTIAIWMPEPTGHTFEFGDRYRFRNEETRGD 497  
 QY 450 FYPPSLQDWGIPPEVSGALPRPDGS-IIFPRDRYWRMLDQAKLQATTSGMATELPM 508  
 Db 498 PGYPRKPIVWQIIPASPKGAFSLSDAATYTFYKGTIKWKFQNERLMEBGPYSILRDFM 557  
 QY 509 GCM-HANGS 516  
 Db 558 GQQRHVEBG 566

Search completed: June 13, 2005, 18:03:21  
 Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 14, 2005, 07:28:01 / Search time 264 Seconds

(without alignments)  
3222.969 Million cell updates/sec

Title: US-10-791-980-6  
2834  
Sequence: 1 MVARGLLRALLQLLWGL.....WATLPMWGMCHANGSALF 520

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODL=Frame+P2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO/epool/US10791980/runat\_13062005\_104719\_19951/app\_query.fasta\_1.711  
-DB=Issued\_Patents\_NA -QFMT=fasta -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITs=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPM=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10791980 @CGN 1.1 177 @runat\_13062005\_104719\_19951 -NCPU=6 -ICPU=3  
-NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length DB | ID | Description       |
|------------|--------|-------------|-----------|----|-------------------|
| 1          | 2834   | 100.0       | 1597      | 4  | US-09-862-631-3   |
| 2          | 2792   | 98.5        | 1560      | 4  | US-09-950-510-3   |
| 3          | 2792   | 98.5        | 2527      | 4  | US-09-950-510-1   |
| 4          | 2764   | 97.5        | 2275      | 3  | US-09-391-104-1   |
| 5          | 2324.5 | 82.0        | 1335      | 4  | US-09-950-510-13  |
| 6          | 2324.5 | 82.0        | 2310      | 4  | US-09-950-510-11  |
| 7          | 602    | 21.2        | 3530      | 3  | US-08-704-711A-10 |
| 8          | 602    | 21.2        | 3530      | 3  | US-09-521-220-10  |
| 9          | 575    | 20.3        | 1524      | 4  | US-09-171-545-7   |
| 10         | 575    | 20.3        | 2264      | 4  | US-09-171-545-22  |
| 11         | 560    | 19.8        | 2049      | 4  | US-09-171-545-23  |
| 12         | 555    | 19.6        | 3437      | 3  | US-08-704-711A-9  |

|    |       |      |      |   |                    |                    |
|----|-------|------|------|---|--------------------|--------------------|
| 13 | 555   | 19.6 | 3437 | 3 | US-09-521-220-9    | Sequence 9, Appli  |
| 14 | 555   | 19.6 | 3437 | 3 | US-09-919-497-34   | Sequence 34, Appli |
| 15 | 554   | 19.5 | 3403 | 4 | US-08-448-489-2    | Sequence 2, Appli  |
| 16 | 554   | 19.5 | 3403 | 4 | US-09-689-730-2    | Sequence 2, Appli  |
| 17 | 551   | 19.4 | 1923 | 3 | US-09-294-841-1    | Sequence 1, Appli  |
| 18 | 550   | 19.4 | 1531 | 4 | US-09-171-545-8    | Sequence 8, Appli  |
| 19 | 538.5 | 19.0 | 2116 | 3 | US-09-000-041A-1   | Sequence 1, Appli  |
| 20 | 538.5 | 19.0 | 2116 | 3 | US-09-734-002-1    | Sequence 1, Appli  |
| 21 | 515   | 18.2 | 2666 | 4 | US-09-949-016-2392 | Sequence 2392, Ap  |
| 22 | 515   | 18.2 | 2666 | 4 | US-09-949-016-2393 | Sequence 2393, Ap  |
| 23 | 515   | 18.2 | 3691 | 3 | US-09-211-704A-3   | Sequence 3, Appli  |
| 24 | 513.5 | 18.1 | 3456 | 3 | US-08-704-711A-8   | Sequence 8, Appli  |
| 25 | 513.5 | 18.1 | 3456 | 3 | US-09-521-220-8    | Sequence 8, Appli  |
| 26 | 506.5 | 17.9 | 3695 | 3 | US-09-211-704A-1   | Sequence 1, Appli  |
| 27 | 492.5 | 17.4 | 2177 | 4 | US-09-919-039-124  | Sequence 124, App  |
| 28 | 478   | 16.9 | 1233 | 4 | US-09-171-545-9    | Sequence 9, Appli  |
| 29 | 478   | 16.9 | 1233 | 4 | US-09-171-545-10   | Sequence 10, Appli |
| 30 | 476.5 | 16.8 | 2248 | 4 | US-09-949-016-4937 | Sequence 4937, Ap  |
| 31 | 476.5 | 16.8 | 2247 | 4 | US-09-023-655-1410 | Sequence 1410, Ap  |
| 32 | 476.5 | 16.8 | 2256 | 1 | US-07-794-393-1    | Sequence 1, Appli  |
| 33 | 476.5 | 16.8 | 2256 | 1 | US-08-001-711-1    | Sequence 1, Appli  |
| 34 | 473   | 16.7 | 2260 | 1 | US-07-794-393-3    | Sequence 3, Appli  |
| 35 | 473   | 16.7 | 2260 | 1 | US-08-001-711-3    | Sequence 3, Appli  |
| 36 | 456   | 16.1 | 1257 | 4 | US-09-171-545-11   | Sequence 11, Appli |
| 37 | 456   | 16.1 | 1272 | 4 | US-09-171-545-12   | Sequence 12, Appli |
| 38 | 427.5 | 15.1 | 1521 | 4 | US-08-994-689C-18  | Sequence 18, Appli |
| 39 | 427.5 | 15.1 | 2792 | 4 | US-08-994-689C-8   | Sequence 8, Appli  |
| 40 | 426.5 | 15.0 | 2717 | 4 | US-09-949-016-4689 | Sequence 4689, Ap  |
| 41 | 426.5 | 15.0 | 1717 | 1 | US-08-228-515A-12  | Sequence 12, Appli |
| 42 | 425.5 | 15.0 | 1717 | 1 | US-08-645-865-1264 | Sequence 1264, Ap  |
| 43 | 403   | 14.2 | 1970 | 4 | US-09-023-655-1264 | Sequence 1264, Ap  |
| 44 | 403   | 14.2 | 1970 | 4 | US-09-949-016-5004 | Sequence 5004, Ap  |
| 45 | 403   | 14.2 | 1973 | 4 | US-09-949-016-352  | Sequence 352, App  |

## ALIGNMENTS

RESULT 1  
US-09-862-631-3  
; Sequence 3, Application US/09862631  
; Parent No. 6734005  
; GENERAL INFORMATION:  
; APPLICANT: Holmgren, Erik  
; APPLICANT: Kihlman, Mats  
; APPLICANT: Wood, Tim  
; APPLICANT: Ekblom, Jonas  
; TITLE OF INVENTION: No. 6734005el Matrix Metalloproteinases  
; FILE REFERENCE: 00014regUS  
; CURRENT APPLICATION NUMBER: US/09/862, 631  
; CURRENT FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 206119  
; PRIOR FILING DATE: 2000-05-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 1597  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-862-631-3

Alignment Scores:  
Pred. No.: 4.9e-226  
Score: 2834.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Gaps: 0

US-10-791-980-6 (1-520) x US-09-862-631-3 (1-1597)

QY 1 MetValAlaArgValGlyValLeuLeuArgAlaLeuGlnLeuLeuTTrGlyHisIleu 20  
DB 35 ATGTTCGGCGCGCTCTGCTCTGCTCTGCGCGCGCTGCGAGCTGCTACTGTGTGGGCGACCTG 94



```

QY      81 A|aThrLeuA|g|lmeTh|ArpProA|g|y|Val|Th|AspTh|AsnSer|y|Ala| 100
DB      241 GCCACCTCGGCAGATGACTGCTCCCGCTGGCGGATTACAGATCAACAGTTATGGC 300
QY      101 A|ATP|Ala|G|U|A|G|T|T|E|Ser|A|Leu|Phe|Ala|Arg|Th|I|y|Met|Arg|G|y| 120
DB      301 GCTGGGCTGAGAGGATGATGATCTTGTGTTGCTAGACACCGGACCAAAATGAGGCTTAAG 360
QY      121 L|y|A|S|Phe|Ala|y|A|S|G|l|n|g|y|A|S|y|P|y|L|y|G|l|n|I|e|u|S|e|r|y|A| 140
DB      361 AAACGTTTGCANAGCAAGGATACAAATGATACACAGACACTTCTCTACGCTGATG 420
QY      141 AAsnTPProG|l|u|n|I|e|u|-Arg|Ser|Arg|l|n|Phe|G|y|Ala|ProCy|Ala|Pro|Pro|Ser|S|e| 160
DB      421 AACTGGCTCGAGATCTGCGCGAGCGGACGTTCCGGGGCGCGCTGGCGCGCTTCCAG 480
QY      160 T|C|y|G|G|y|Ala|Th|S|e|r|G|l|n|A|T|P|S|e|r|S|e|r|G|y|Ala|P|ro|G|l|n|A|P|ro|Leu|Th| 180
DB      481 TTGTGGAGCAAGCTCTCAGCGCTGAGATTCTGGGAGCGCCACAGCCACAGCGCCCGCTGAC 540
QY      180 r|S|e|r|G|y|S|e|P|ro|S|e|r|S|e|r|S|e|r|y|Th|Th|Th|Met|G|y|T|P|Ala|Met|Pro|Leu|Met|Al| 200
DB      541 ATCCGAGTCACTTCTTCCAGGGAGCACACAGATGGCTGGGCAATGCTTTGATGGC 600
QY      200 a|G|l|n|g|y|Ala|P|ro|T|P|Arg|Th|P|ro|Phe|u|P|ro|A|T|Arg|G|y|G|l|u|Ala|H|I|S|Phe|A|P|G|l| 220
DB      601 CCAGGGGGGGCCCTGGCGCACGCTTC-C-CTGGCCCGCGCGGAGAGCCGACCTTGACCA 659
QY      220 n|A|S|P|G|l|u|A|G|T|P|S|e|r|L|e|u|S|e|r|A|G|A|R|G|y|A|T|G|A|n|L|e|u|Phe|Val|I|e|u|Ala|H| 240
DB      660 AATATGAGCCTGGTCCCTAGCGCGCGCGCGCGCCCAACCTGTTGTGTGCTGCGCGCA 719
QY      240 a|G|l|u|I|e|G|y|H|e|Th|L|e|u|G|l|y|L|e|u|Th|H|I|S|e|r|P|ro|A|P|ro|A|G|l|e|u|Met|Ala|P| 260
DB      720 CGAGATCGGTCACACGCTTGCTGCTCACCACTGCGCGCGCGCGCGCTCATGGCGGC 779
QY      260 c|T|y|T|y|L|y|A|T|G|A|T|G|L|y|A|S|A|P|Ala|L|e|u|L|e|u|S|e|r|T|P|A|S|P|A|S|P|Val|I|e|u|Ala|G|l| 280
DB      780 CTAACATCAAGAGGCTGGCGCGCGCGCTGCTCACTGGAGAGAGTCTGGCCCTGCA 839
QY      280 n|S|e|r|L|e|u|T|y|G|y|L|y|P|ro|L|e|u|G|y|G|y|S|e|r|Val|Ala|Val|G|l|n|L|e|u|P|ro|G|l|y|L|y|L|e|u|H| 300
DB      840 GAGCCCTGATGGGAGAACCCCTAGGGGCTCAGTGGCGCTGCAAGCTCCACGAAACGCTGTT 899
QY      300 e|H|I|S|e|r|P|he|G|l|u|Th|T|P|A|S|P|S|e|r|y|S|e|r|P|ro|G|l|n|g|y|A|G|A|T|G|P|ro|G|l|u|Th|G|l|n|G|l| 320
DB      900 CACTGACTTTGAGACTGGGACTCTTACAGCCCCCAAGAGGCGCGCTGAACCGCAGG 959
QY      320 y|P|ro|L|y|T|y|C|y|H|I|S|e|S|e|S|e|Phe|A|P|Ala|I|e|Th|r|Val|A|S|P|A|S|G|l|n|G|l|n|L|e|u|T|y| 340
DB      960 CCTCAATAACTGCGCACTCTTCTCTGATGCACTGTAAGACAGGCAACAGAACTGTA 1019
QY      340 r|I|e|Phe|y|G|y|S|e|H|I|S|e|Phe|T|P|G|l|u|Ala|Ala|A|S|P|G|y|A|S|e|r|L|y|P|ro|A| 360
DB      1020 CATTTTAAAGGAGGACATTTCTTGAGGAGCTGAGAGCTGATGAGCAAGTCTCAGAGCCCG 1079
QY      360 g|P|ro|L|e|u|G|l|n|g|y|T|P|Val|G|l|y|L|e|u|P|ro|A|S|n|I|e|G|l|u|Ala|Ala|Val|S|e|r|S|e| 380
DB      1080 TCACATGCAAGAAAGATGGAGTGGGCTGCCCCCAACATTGAGGCTGCGCAGTGCATT 1139
QY      380 u|A|S|P|A|S|P|G|y|A|S|P|Phe|T|y|Phe|L|y|S|G|y|A|T|G|C|y|S|T|P|A|S|P|A|S|P|A|S|P|G|y|P|ro|L|y| 400
DB      1140 GAATATGAGATTTCTACTTCTTCAAGGGGCTCAATGCTGAGGTTCCGGGGCCCA 1199
QY      400 a|P|ro|Val|T|P|G|y|L|y|L|e|u|P|ro|G|l|n|L|e|u|Cy|A|T|G|Ala|G|l|y|L|y|L|e|u|P|ro|A|S|H|I|S|P|A|S|P|Al| 420
DB      1200 GCCAGTGTGGGCTCCCAAGCTGTGCGGGGCAAGGGGCGCTGCCCCGCACTTCAAGCC 1259
QY      420 A|A|T|P|Ala|G|l|u|A|G|T|T|E|S|e|r|A|Phe|Phe|Ala|Arg|Th|I|y|Met|Arg|G|y|S| 440
DB      1260 GCGCTCTTCTTCTCTCTCTGCGCGCTCATCTTCAAGAGGTCGCGCTACTACGT 1319

```

```

QY      440 l|Leu|A|A|A|A|G|y|G|y|L|e|u|G|l|n|Val|G|l|u|P|ro|T|y|T|P|P|ro|A|S|e|r|L|e|u|G|l|A|S|P|T|P|G|l| 460
DB      1320 GCTGGCCCCGAGGGGACCTGCAAGTAGAGCCCTACTACACCCCAAGCTGTCAGACAGCGGG 1379
QY      460 y|G|y|L|y|P|ro|G|l|u|G|l|n|S|e|r|G|y|Ala|L|e|u|P|ro|A|S|P|A|S|P|G|y|S|e|r|I|e|I|Phe|H| 480
DB      1380 AGGCAATCCCTGAGAGGATGAGCGGCGCTGCGAGGCGCGATGGCTCATCATCTTCTT 1439
QY      480 e|A|R|S|P|A|S|P|A|S|P|G|y|T|P|P|A|S|P|G|y|Ala|S|e|r|G|l|n|A|Th|Th|S|e|r|G|y|A| 500
DB      1440 CCGAGATGACCCGCTACTGCGCGCTGACCAAGGCCAAACTGACAGGCAACCACTCGGGCGG 1499
QY      500 g|T|P|Ala|Th|G|l|u|L|e|u|P|ro|T|P|Met|G|y|C|y|S|T|P|H|I|S|A|A|S|n|S|e|r|G|y|S|e|r|Ala|L|e|u|H| 520
DB      1500 CTGGGCCACCGAGCTGCGCTGTGATGGGCTGCTGGCATGCGCAACTCGGGAGCGCCCTGTT 1559
QY      520 e 520
DB      1560 C 1560

RESULT 3
US-09-950-510-1
; Sequence 1, Application US/09950510
; Patent No. 6740514
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory
; TITLE OF INVENTION: 46798, A No. 6740514el Human Matrix Metalloproteinase And Uses Th
; FILE REFERENCE: 10147-45ul
; CURRENT APPLICATION NUMBER: US/09/950,510
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/251,156
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patencin version 3.0
; SEQ ID NO 1
; LENGTH: 2527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-950-510-1

Alignment Scores:
Pred. No.: 2,94e-222 Length: 2527
Score: 2792.00 Matches: 519
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 98.52% Indels: 2
DB: 4 Gaps: 0

US-10-791-980-6 (1-520) x US-09-950-510-1 (1-2527)
QY      1 MetValAlaArgValG|L|y|L|e|u|L|e|u|A|L|e|u|G|l|n|L|e|u|L|e|u|T|P|G|l|H|I|S|L|e|u| 20
DB      300 ATGTGTGCGCGGCTGCTCTTCTGCTGCGCGCTTGCAGCTCTACTGTGGGGCCACCTG 359
QY      21 AAsPAlaIeArgProAlaG|U|A|G|y|G|y|G|l|n|L|e|u|A|T|G|y|G|l|u|AlaG|l|u|A|A|Phe|L|e|u| 40
DB      360 GAGCGCCAGCCCGCGGAGGCGGAGGCGCAGAGCTGCCGAAGAGGCGGAGGCAATTCCTA 419
QY      41 G|U|L|y|T|y|G|y|T|y|L|e|u|A|n|G|l|n|Val|P|ro|L|y|A|P|ro|Th|S|e|r|Th|Arg|Phe|S|e|r| 60
DB      420 GAGAGTACCGGATTCATGATGACAGGTCGCCCAAGCTCCCACTCCCACTGATTCAGC 479
QY      61 AAsPAlaIeArgAlaPheG|l|n|T|P|Val|S|e|r|L|e|u|P|ro|Val|S|e|r|G|y|Val|L|e|u|A|S|P|A|S|P| 80
DB      480 GATGCCATCAGAGCGTTTCAAGTGGATGCCAGCTACTCTGACGCGCGCTGTGACCC 539
QY      81 A|A|T|P|Ala|G|l|u|A|G|T|T|E|S|e|r|A|Phe|Phe|Ala|Arg|Th|I|y|Met|Arg|G|y|S| 100
DB      540 GCCACCTCGGCAGATGACTGCTCCCGCTGCGGGTTACAGATACCAACAGTTATGGC 599
QY      101 A|A|T|P|Ala|G|l|u|A|G|T|T|E|S|e|r|A|Phe|Phe|Ala|Arg|Th|I|y|Met|Arg|G|y|S| 120
DB      600 GCTGGGCTGAGAGATGATGATCTTGTGTTGCTAGACACCGGACCAAAATGAGGCTTAAG 659

```



Db 406 AAACGCTTTGCAAGCAAGGTGACAAATGCTACAGCAGCACTCTCTTACCGCGCTGGTG 465  
 Qy 141 AentPrProGluHhIeUeu-ArgSerArgGlnPheGlyAlaProCyAlaProProSerSe 160  
 Db 466 AACTGGCCCTGAGCATCTTCCGAGCGGAGATTCCGGGGGCGCGTGGCCGCCCTTCCAG 525  
 Qy 160 rCyGgAlaIaThrSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuTh 180  
 Db 526 TTGTGAGACAGCATCTCAAGCGCTGAGATTCTGGAGAGCCCAAGCCAGAGCCCGCTGAC 585  
 Qy 180 rSerGlySerProSerSerSerGlyThrThrThrMetGlyTrpAlaMetProLeuMetAl 200  
 Db 566 ATCCGGCTCACCCTTCTTCCAAAGGGAGCCAAAGATGGGCTGGGCAATCCCTTTGATGGC 645  
 Qy 200 aGlnGlyAlaProTrpArgTrpPheLeuProArgArgGlyGlyAlaHhIePheApsl 220  
 Db 646 CCAAGGGGGGGCCCTGGCGCAAGCCCTTC-CTGCCCGCCCGCGGGAATTTTACTTCAGCA 704  
 Qy 220 rApslGlyArgTrpSerSerSerArgArgArgGlyArgAenLeuPheValIaLeuAlaHh 240  
 Db 705 AGATAGAGCGCTGCTCCCTGAGCGCGCGCGGCGCAACTGTTCGTGTGCTGGCGCA 764  
 Qy 240 aGluIleGlyHhIeThrLeuGlyLeuThrHhIeSerProAlaProArgAlaLeuMetAlaP 260  
 Db 765 CGAGATCCGGTCAACGCTTGGCTTCCACCTCGCCCGCGCGCGCTCATGCGCGC 824  
 Qy 260 CTATTrpIleArgArgLeuGlyArgApsAlaLeuLeuSerTrpApsApsValIaValG 280  
 Db 825 CTACTCAAGAGAGCGCTGGCGCGAGCGCGCTGCTCACTGGAGAGAGCTGCTGGCGTGA 884  
 Qy 280 nSerLeuTrpGlyIleArgProLeuGlyGlySerValAlaValGlnLeuProGlyIleVal 300  
 Db 885 GAGCGCTGTATGGGAAGCCCTTACGGGGGCTCAGTGGCGCTCCAGCTCCAGAAAGCTGT 944  
 Qy 300 eThrApsPheGlnIleThrTrpApsSerTrpSerProGlnGlyArgArgProGlnIleG 320  
 Db 945 CACTACCTTTGAGAGCTGGGAGCTCTTACAGCCCCCAAGAGAGCGCGCTGAAACGAGG 1004  
 Qy 320 rProIleTrpCyHhIeSerSerPheApsAlaIleThrValApsArgGlnGlnLeuTrp 340  
 Db 1005 CCTCAATATCTGCCACTCTTCTTGCATGCACTCACTGAGACAGGCAACAGCACTGTA 1064  
 Qy 340 rIlePheIleGlySerHhIePheTrpGlyValAlaApsGlyApsValSerGlnProAr 360  
 Db 1065 CATTTTAAAGGAGGACATTTCTGGGAGGTGGCACTGATGGCAAGTCTCAGAGACCGG 1124  
 Qy 360 gProLeuGlnGlyIleArgTrpValGlyLeuProProAenIleGlyAlaAlaValSerIe 380  
 Db 1125 TCCTACTGCAAGAAAGATGGGTGGGTGGCTGCCCAACATTTGAGGCTGGCAGTGTCA 1184  
 Qy 380 uAenApsGlyApsPheTrpPhePheIleGlyIleArgCyHhIeTrpArgPheArgGlyPro 400  
 Db 1185 GAAATATGAGATTTCTACTTCTTCAAAAGGGAGTCACTGAGAGTTTCGGGGGCCCA 1244  
 Qy 400 sProValTrpGlyLeuProGlnLeuCyApsArgAlaGlyIleGlyLeuProArgHhIeProA 420  
 Db 1245 GCCAGTGGGGGTCTCCCAAGCTGGGCGAGGGGCGTGGCCCGCATCTGAGAGC 1304  
 Qy 420 aAlaLeuPhePheProProLeuApsArgArgLeuIleLeuPheIleGlyAlaArgTrpVal 440  
 Db 1305 CGCCCTCTTCTTCCCTCTGCGCGCTCACTCTTCAAGGGGCGCGCTACTACTACT 1364  
 Qy 440 lLeuAlaArgGlyIleGlyLeuGlnValGlnProTrpTrpProArgSerLeuGlnApsTrp 460  
 Db 1365 GCTGGCGCGAGGGAGCTCAAGTGGAGCCCTACTACCCCAAGTCTTCCAGAGATGGGG 1424  
 Qy 460 rGlyIleProGlnGlyValSerGlyAlaLeuProArgProApsGlySerIleIlePheHh 480  
 Db 1425 AGGCATCCCTGAGAGGTCAAGCGCGCGCTTCCGAGGCGCGAGTGGCTCATCTTCTT 1484  
 Qy 480 eArgApsApsArgTrpTrpArgLeuApsGlnAlaIleValLeuGlnAlaThrThrSerGlyA 500  
 Db 1485 CCGAGATGACCGCTACTGGCGCTCGAACAGGCGCAAACTGACAGGAAACCACTCGGGCG 1544

Qy 500 gTrpAlaThrGlyLeuProTrpMetGlyCyHhIeHhIeAenSerGlySerAlaLeuHh 520  
 Db 1545 CTGGCGCAACGAGCTGCCCTGTGATGGGCTGTGCGATGCACTCGGGAGCGCGCTGT 1604  
 Qy 520 e 520  
 Db 1605 C 1605  
 RESULT 5  
 US-09-950-510-13  
 ; Sequence 13, Application US/09950510  
 ; Patent No. 6740514  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Curtis, Rory  
 ; TITLE OF INVENTION: 46798, A No. 6740514el Human Matrix Metalloproteinase And Uses Tr  
 ; FILE REFERENCE: 10147-45U1  
 ; CURRENT APPLICATION NUMBER: US/09/950,510  
 ; PRIOR FILING DATE: 2001-09-10  
 ; PRIOR APPLICATION NUMBER: 60/251,156  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 13  
 ; LENGTH: 1335  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-950-510-13  
 Alignment Scores:  
 Pred. No.: 8,42e-184 Length: 1335  
 Score: 2324.50 Matches: 444  
 Percent Similarity: 85.22% Conservative: 0  
 Best Local Similarity: 85.22% Mismatches: 1  
 Query Match: 82.02% Indels: 77  
 DB: 4 Gaps: 1  
 US-10-791-980-6 (1-520) x US-09-950-510-13 (1-1335)  
 Qy 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHhIeU 20  
 Db 1 ATGTGTGGCGCGGTGGCTCTGCTGGCGCGCTGCGAGCTGCTACTGTGGGGCCACCTG 60  
 Qy 21 ApsAlaGlnProAlaGlnArgGlyIleGlnGlnLeuArgGlyValAlaGlnAlaPheLeu 40  
 Db 61 GAGCGCCAGCGCGCGGAGGCGGAGCCAGAGCTGCCGGAAGAGAGCGGAGCATTCCTTA 120  
 Qy 41 GlnIleTrpGlyTrpLeuAenGlnValProIleAlaProThrSerThrArgPheSer 60  
 Db 121 GAGAGTACGATACCTCAATGAACAGGTCCCAAGCTCCCACTCCACTGATTCAGC 180  
 Qy 61 ApsAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuApsArg 80  
 Db 181 GATGCCATCAAGCGCTTCAAGTGGGTGTCCAGCTACTGTCAGGGCGGTGTGACCGC 240  
 Qy 81 AlaThrLeuArgGlnMetThrArgProArgCyGlyValAlaThrApsThrApsSerTrpAla 100  
 Db 241 GCCACCTCGCGCAATGACTGCTCCCGCGGGGTTCAGATACCAACAGTTATGG 300  
 Qy 101 AlaTrpAlaGlnArgIleSerApsLeuPheAlaArgHhIeArgThrIleMetArgArgIle 120  
 Db 301 GCGTGGCGTGAAGGATCACTGCTTTGCTGACACCGGACCAAAATGAGCGGTAA 360  
 Qy 121 LysArgPheAlaIleGlnGlyApsIleTrpTrpIleGlnIleLeuSerTrpArgLeuVal 140  
 Db 361 AAACGCTTGCAAA- 375  
 Qy 141 AentPrProGluHhIeUeuArgSerArgGlnPheGlyAlaProCyAlaProProSerSer 160  
 Db 375 ----- 375  
 Qy 161 CyGlyAlaThrSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuThr 180

```

Db      375 -----
QY      181 SerGlySerProSerSerlySGlyThrThrMetGlyTyrPalamcProLeuMetAla 200
Db      375 -----
QY      201 Gln-GlyAlaProTTPArGThrProPheLeuProArGArglyGlyGlnAlaHisPheArgI 220
Db      376 CAAGGGGGCGCCCTGGCGCAGCCCTTC-CTGCCCCGGCGCCAGAGCGCACTTCACCA 434
QY      220 NAARGIUAArgTTPSerLeuSerArGArGlyArGAsnLeuPheValIleuAlaHis 240
Db      435 AGATGAGCCCTGGTCCCTGAGCCGCCCGCGGGCGCAACCTGTCTGTGTGGCGCA 494
QY      240 eGluIleGlyHisGThrLeuGlyLeuThrHisSerProAlaProArGAlaLeuMetAlaP 260
Db      495 CGAGATCGGTCAACAGCTTGGCTTCACCCACTCGCCCGCGCCGCGCTCAGTGGCGCC 554
QY      260 CTYrTyrIlysaArgLeuGlyArGAspAlaLeuLeuSerTTPAspAspValIleuAla 280
Db      555 CTACTACAAAGAGCTGGGGCGCGAGCGCTGTCTACCTGGAGCAAGTGTGGCTGCCTCA 614
QY      280 nSerLeuTyrGlySerProLeuGlyGlySerValAlaValGlnLeuProGlyIlyLeuP 300
Db      615 GAGCTGTATGGAGAGCCCTAGAGGGGCTCAGTGGCGCTCCAGCTCCAGAGAGCTGTT 674
QY      300 eThAspPheGlyThrTTPAspSerTyrSerProGlnGlyArGArgProGlnThrGlnI 320
Db      675 CACTGACTTTGAGACCTGGGACTCTCAACGCCCCCAAGAGAGGGCGCCCTGAAACGCGGG 724
QY      320 YProLyTyTyCyHisSerSerPheAspAlaIleThrValAspArgGlnGlnLeuTyr 340
Db      735 CCTAAATATGCGCACTTCTCTGATGCGCATCTAGTACAGGCAACGCAACTGTGA 794
QY      340 rIlePheLySGlySerHisPheTyrGlnValAlaAlaAspGlyAsnValSerGluProAr 360
Db      795 CATTTTAAAGGAGCGCAATTTCTGGAGCTGCGACGTGAGTGGCACTTCACAGGCCCG 854
QY      360 gProLeuGlnGlnUaSGTTPValGlyLeuProProAsnIleGlnAlaAlaValSerIle 380
Db      855 TCACATGCGAGAAAGATGGAGTGGGCTGCCCCCAACATTCAGAGCTCGCGGCACTGTCA 914
QY      380 uAsnAspGlyAspPheTyrPhePheLySGlyIlyArGlyCySerTPArGArgGlyProLy 400
Db      915 GAATGATGAGATTTCTACTTCTTCMAAGGGGCTCACTGTGAGGTTCCGGGGCCCCAA 974
QY      400 sProValTyrGlyLeuProGlnLeuCyAsnAlaGlyGlyLeuProArGHisProAspAl 420
Db      975 GCCAGTGGGGGTCTCCCAACGCTGTGCGGGAGGGGGCGTGGCCCGCATCTGACGC 1034
QY      420 aAlaLeuPhePheProProLeuArGArgLeuIleLeuPheLySGlyAlaArgTyrTyrVa 440
Db      1035 CGCCCTCTTTCTCTCTCTCTGCGCGCTCATCTCTTCMAAGGGTGGCCGCTACTACGT 1094
QY      440 lIleuAlaArGlyGlyIlyLeuGlnValGlnProCyrTyrProArGSerLeuGlnAspTyrG 460
Db      1095 GCTGGCGCCGAGGGGAGCTCAAGTGAAGCCCTACTACCCCGAAGTCTGCGAGGACTGGGG 1154
QY      460 yGlyIleProGlnGlnValSerGlyAlaLeuProArGProAspGlySerIleIlePheP 480
Db      1155 AGGCATCTCTGAGAGGTCAAGCGCGCTCGCCAGAGCCGAGTGGCTCCATCACTTCTTT 1214
QY      480 eArGAspAspArGlyTyrTPArGLeuAspGlnAlaIlyLeuGlnAlaThrThrSerGlyAr 500
Db      1215 CCGAGATGACCGCTACTGCGCCCTCGACACAGGCCAACTGCAAGCAACCACTCGGGCGG 1274
QY      500 GTTPAlaThrGlnLeuProTTPMetGlyCySerTPHisAlaAsnSerGlySerAlaLeuP 520
Db      1275 CTGGGCCACCGAGCTGCTGAGTGGGCTGTGCGCATGCCAACTCGGGAGGCGCTGTGT 1334
QY      520 e 520
Db      1335 C 1335

```

```

RESULT 6
US-09-950-510-11
/ Sequence 11, Application US/09950510
/ Patent No. 6740514
/ GENERAL INFORMATION:
/ APPLICANT: CurieB, Roy
/ TITLE OF INVENTION: 46798, A No. 6740514el Human Matrix Metalloproteinase And Uses T
/ FILE REFERENCE: 10147-45U1
/ CURRENT APPLICATION NUMBER: US/09/950,510
/ PRIOR APPLICATION NUMBER: 2001-09-10
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patent version 3.0
/ SEQ ID NO 11
/ LENGTH: 2310
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-950-510-11

Alignment Scores:
Pred. No.: 1,86e-183 Length: 2210
Score: 2324.50 Matches: 44
Percent Similarity: 85.22% Conservative: 0
Best Local Similarity: 85.22% Mismatches: 1
Query Match: 82.02% Indels: 77
DB: Gaps: 1

US-10-791-980-6 (1-520) x US-09-950-510-11 (1-2310)
QY      1 MetValAlaArgValGlyLeuLeuLeuArGAlaLeuGlnLeuLeuTTPGlyHisLeu 20
Db      317 ATGATGCCCGGCTGCGCTCTGCTGGCGCCCTCGACAGTCTACTGTGGGGCGCACTG 376
QY      21 AspAlaGlnProAlaGlyIlyGlyGlnGlnLeuArgIlySGlyAlaGlnAlaPheLeu 40
Db      377 GAGCGCCAGCCCGCGAGCGGAGGCCAGGAGCTGGCGCAAGAGGGCGAGGCACTTCTTA 436
QY      41 GlnUlyTyTyTyTyLeuAsnGlnGlnValProLyValProPheSerThrArGPheser 60
Db      437 GAAAGTACGGAATCTCTCAATGAACAGTCCCAAGCTCCCACTCCACTGATTCAAC 496
QY      61 AspAlaIleAspAlaPheGlnTTPValSerGlnLeuProValSerGlyValIleuAspArg 80
Db      497 GATGCCATCAAGCGTTTCACTGAGTGTGTCCCACTACTGTGACGGCGTTTGACCGC 556
QY      81 AlaThrLeuArGlnMetThrArGProArGlyGlyValThrAspThrAsnSerTyrAla 100
Db      557 GCCACCTGGGCCAGATGACTGCTCCCGCTGGGGTTACAGATACCAACAGTTATCGG 616
QY      101 AlaTPAlaGlnArgIleSerAspLeuPheAlaArgHisArgThrIlyMetArGArglyS 120
Db      617 GCTGGGCTGAGAGATCACTGACTTGTGTCTACACCGGAGCAAAATGAGGCGTTAG 676
QY      121 IlyAspPheAlaIlySGlnGlyAsnLySTyrIlySGlnHisLeuSerTyrArGLeuVal 140
Db      677 AAACGCTTGGCAAG-----
QY      141 AsnTTPProGlnHisLeuArGSerArGlnPheGlyAlaProCyValaProProSerSer 160
Db      691 -----
QY      161 CySGlyAlaThrSerGlnArGTrpSerSerGlyArGProGlnProGlnAlaProLeuThr 180
Db      691 -----
QY      181 SerGlySerProSerSerlySGlyThrThrMetGlyTyrPalamcProLeuMetAla 200
Db      691 -----
QY      201 Gln-GlyAlaProTTPArGThrProPheLeuProArGArglyGlyGlnAlaHisPheArgI 220

```



|    |             |   |      |
|----|-------------|---|------|
| Db | 692         | CAAGGGGGGCGCCCTGCGCAGCGCTTC-CTGGCCCGCGCGGAGGAAGCGCACTTGACCA     | 75   |
| Qy | 220         | naapgluarrgttppserleuSerAArgArgrglYargAenluPheValValleuAlaI     | 240  |
| Db | 751         | AGATAGGCGCTGGTCTCTGAGCGCGCGCGCGGAGCAACTGTCGTGGTCTGGCGCA         | 810  |
| Qy | 240         | agluilegIyIethrleuGlyleuThrHisSerProAlaProArgAlaLeuMecAlaPr     | 260  |
| Db | 811         | CGAGATCCGCTCACAGCGTTGGCTTACCACCTCGCCCGCGCGCGCGCTCATGGCGCC       | 870  |
| Qy | 260         | oTyTylYargleuGlyArgAspAlaLeuIleuSerTTPaAspAspValleuAlaIGI       | 280  |
| Db | 871         | CTACTACAAGAGGCTGGGCGCGCGAGCGCGCTGCTCAGCTGGAGCAAGCTGCGCGTGA      | 930  |
| Qy | 280         | nSerleuTyrglyYlvsProleuGlyGlySerValAlaValIGInleuProGlyYlyleuPh  | 300  |
| Db | 931         | GAGCTGTATGGGAAGCCCTTAGGGGGCTCAGATGGCCGCTCCAGCTCCAGGAAGCTGTT     | 990  |
| Qy | 300         | eThraaphegluThrTPaasPserTysSerProGInglyArgArgProGlyluthrgInGI   | 320  |
| Db | 991         | CACGTACCTTTGAGACCTGGGACTCTTACAGCCCCCAAGGAAGCGCCCTGMAACGAGGG     | 1050 |
| Qy | 320         | yProlystTyCyvHisSerSerPheAspAlaIethrValaAspArgInGInleuTy        | 340  |
| Db | 1051        | CCCTAAATACCTCCCACTCTTCTTGATGCATCACTGATGACAGGCAACAGCAACTGTAT     | 1110 |
| Qy | 340         | rIlePheIyagIySerHisPheTTPGIValAlaAlaAspGlyAsenValSerGIuProAr    | 360  |
| Db | 1111        | CATTTTAAAGGAGCCCAATTTCTGGAGGTGCAGCTGATGGCAACGTCTCAGAGCCCCG      | 1170 |
| Qy | 360         | gProleuGIngluAsgTTPValIGlyleuProProAsnIleGluAlaAlaValSerIe      | 380  |
| Db | 1171        | TCCACTGCAGAAABATGGGTGGGGCTGCCCCCAACATTAGCGTGGCGGAGTGCATT        | 1230 |
| Qy | 380         | uaenapglYasPheTyrrPhePheIyagIyIArgCystrParapheArGlyPProly       | 400  |
| Db | 1231        | GAATGATGAGATTTCTACTTCTTCAAGAGGGGTGCAATGCTGAGGGTTCGGGGCCCCAA     | 1290 |
| Qy | 400         | aProValITPGLyleuProGInleuCyArgrglAGlyGlyleuProArghIsProAspAl    | 420  |
| Db | 1291        | GCCAGTGGGGGTCTCCACACACTGTGCCGGGAGGGGCGCCCGCCATCTGTACCC          | 1350 |
| Qy | 420         | aAlaIeuphePheProProleuArArgIleuIleuPheIyagIyAlaArgTyrrVa        | 440  |
| Db | 1351        | CGCCCTCTTCTTCCCTCTCTGCGCGCTCATCTCTTCAAGGGTGGCCGCTACACGT         | 1410 |
| Qy | 440         | lleuAlaArgGlyGlyleuGInValIGluProTyrrTyrrProArgSerleuGInAspITPGL | 460  |
| Db | 1411        | GCTGGCCCGAGGGGAGCTGCAAGTGGAGGCCCTACTACCCCGAAGCTCAGAGACTGGGG     | 1470 |
| Qy | 460         | yGlyYlIeProGluGluValSerGlyAlaIleuProArgrProAspGlySerIleIlePhePh | 480  |
| Db | 1471        | AGGCATCCCTGAGGAGGTCAAGGGGCGCTGTGCCAGGGCCGATGGTTCATCATCTTCTT     | 1530 |
| Qy | 480         | eArgAspAspArgTyrrTPArgleuAspGlnAlaIySleuGInAlaIthrThrsSerGIYAr  | 500  |
| Db | 1531        | CCGAGATACCCGCTACCTGGCGCTCGAACAGGCCAAACTGCAGGCAACCACTCGGGCGG     | 1590 |
| Qy | 500         | gTTPAlaThrGluIeupProITPMeGlyCYsITPHisAlaAsnSerGlySerAlaIeupH    | 520  |
| Db | 1591        | CTGGGCGCACCGAGCTGCCCTGGATGGGCTGTGGATGCAACTCGGGAGAGCGCCCTGTT     | 1650 |
| Qy | 520 e 520   |   |      |
| Db | 1651 c 1651 |   |      |

RESULT 7  
US-08-704-711A-10  
; Sequence 10, Application US/08704711A  
; Patent No. 6114159  
; GENERAL INFORMATION:  
; APPLICANT: WILL, Horst  
; APPLICANT: HINZMAN, Bernd

```

? TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
?
? NUMBER OF SEQUENCES: 22
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Foley & Lardner
? STREET: 3000 K Street, N.W., Suite 500
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20007-5109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/704,711A
? FILING DATE: 20-Nov-1996
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: WO PCT/DE95/00357
? FILING DATE: 17-Mar-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: DE 4438838.1
? FILING DATE: 21-Oct-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: DE 4409663.1
? FILING DATE: 17-Mar-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: GRANADOS, Patricia D.
? REGISTRATION NUMBER: 33,683
? REFERENCE/DOCKET NUMBER: 26083/124
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)672-5300
? TELEFAX: (202)672-5399
? TELEX: 904136
? INFORMATION FOR SEQ ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3530 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
US-08-704-711A-10

```

Alignment Scores:

| Pred. No.:             | Score: |
|------------------------|--------|
| Percent Similarity:    | 45.12% |
| Best Local Similarity: | 32.50% |
| Query Match:           | 21.24% |
| DB:                    | 3      |

Gaps: 20

US-10-791-980-6 (1-520) x US-08-704-711A-10 (1-3530)

OY    9 LeuaGalaLeugInleuleuLeutTpgIyhiSeuaApalaglInProalagluArgGly 28  
     ||||| :|||||:::||| |  
Db    133 CTGCCTCCCGCCTGTCTTCGGTGCTGCCAGATCCCTGACGTTGGCAAGCG----- 183  
OY    29 GLyglnGlueLuarglyseglualaglualaPhetelengluystryrgllyrrtleuanglu 48  
     ::::::: ||||| :|||||:::||| |  
Db    184 GCCGAAGAACC GGAGSTGCATCCGACA GAAC TGCGCGGCTTTAA TNGCTAAC TCCTCAC 243  
OY    49 GlNvalPrrolyaLaPthrThserThrArg-----PhaseAspaAlale 63  
     ||||| :|||||:::||| |  
Db    244 -----CCNAGCCGCATTATGTCACCATGCGCTCCC CGCCAGACTTGGCTCGGCCCTT 297  
OY    64 ArgAlaPhnegIntTPValSerInserrleuProvallserGIyaLLeuaAParAglaTFrLeu 83  
     :|||: :|||: :|||: :|||: :|||:  
Db    298 GCAGAGATCGACGGCTTYTAACGAGATCCAGTCAACGGGTGTGCTGCAGCAAAGACCAAG 357  
OY    84 ArsglmethrArgrproargCyseglyva]ThrAprHraenserryralaalatrpala 103  
     ||||| :|||||:::||| |  
Db    358 GAQTGATGAAGGGCCCCGGCTGTGAGGAGCCAGC-----CA GTTGGG 402  
                                :::|

QY 104 GluArg1LeuSerAspLeuPheAlaArgHisArgThrLysMetArgAluLysIleArgPhe 123  
 Db 403 GTCAGAGTGA-----GCCAACCTGGCGGCGCTCGGAGAGCGCTAC 444  
 QY 124 AlaLysGlnGlyAsnLysTrpTyrLysGlnIleLeuSerTyrArgLeuValAsnTrpPro 143  
 Db 445 GCCCTCACCGGAGGAGAGTGAACAACAACCATCTGACCTTTAGCATCCAGACTACACG 504  
 QY 144 GluHisLeu-ArgSerArgGlnPheGlyValaProCysAlaProProSerSerCysGlyAla 163  
 Db 505 GAGAAGTTGGCTGTGATACCATCATGAGCGGCTGCCAGGCGCTTCGCGTGTGGAG 564  
 QY 163 arHisSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuThrSerGlyse 183  
 Db 565 CAGCGCAGCGCCCTGCTCTCCAGAGG-----TGCCCTTAGAGACATCCG 612  
 QY 183 rProSerLysGlyThrThrMetGlyTrpAlaMetPro----- 197  
 Db 613 CTGCGCGCAGAGAGAGCGGACATCATGTGACTCTTGCTTGCTTCACAGCGGAC 672  
 QY 198 -----LeuMetAlaGlnGlyAlaProTrpArgThrProPheLeuPro----- 211  
 Db 673 AGCTGCGCTTGATGAGCAGCGGTGCTTCTGCGCCACGCTAT-TTCCCTGGCCCGG 731  
 QY 212 -ArgArgGlyGlnAlaHisPheAspGlnAspGlnArgTrpSerLeuSerArgArg----- 229  
 Db 732 CCTAGGGCGGGACACCATTTTGAAGCAGATGAGCCCTTGACCTTCCAGCATGACCT 791  
 QY 230 -ArgGlyArgAsnLeuPheValIleuAlaHisGluIleGlyHisThrLeuGlyLeuTh 249  
 Db 792 GCATGGAACAACCTCTCTGCTGAGCAGTGAAGCTGGGCAAGCGCTGGGGCTGA 851  
 QY 249 rHisSerProAlaProArgAlaLeuMetAlaProTyrTyrLysArgLeuGlyArgAspAla 269  
 Db 852 GCACTCCAGCAACCCCATGCTCATGCGCGCTTCTACCACTGAGAGAGAGCTTGACAA 911  
 QY 269 Aleu-----LeuSerTrpAspAspValIleuAlaGlnSerLeuTyrGlyLysProLeuG 288  
 Db 912 CTTCAAGCTGCCGAGAGACATCTCGGTGGCATCCAGCAGCTTACGGTACCCAGACGG 971  
 QY 288 Y-----GlySerValAlaValGlnLeuProGlyLys-Leu 300  
 Db 972 TCAGCCACAGCCTACCCAGCTCTCCCACTGTGACGCCACGCGCGGCGGCGGCTGA 1031  
 QY 300 heThrAspPheGlnThrTrpAspSerTyrSerProGln-GlyArg----- 314  
 Db 1032 CCAACGCGCGCGCGCGCTCCCAAGCACCACCCCAAGTGGAGAGCCAGCGCGCC 1091  
 QY 315 -----ArgProGlnThrGlnGlyPro 322  
 Db 1092 AAAACCG 1151  
 QY 322 rTyrCysHisSerSerPheAspAlaIleThrValAspArgGlnGlnLeuTyrIlePhe 342  
 Db 1152 CATGTGCGACGGGAGCTTGAACAGTGGCCATGCTTGC---GGGAGATGTTCTGT 1208  
 QY 342 elyGlySerHisPheTrpGluValAlaAlaAspGlyAsnValSerGlu--ProArgPr 361  
 Db 1209 CAAGGCGCGCTGTCTGCGCAGTCCGCGCAACCGGCTCTGAGCAACATCCATGCC 1268  
 QY 361 OLeuGlnGlnuArgTrpValGlyLeuProProAsnIleGluAlaAlaValSerLeuAs 381  
 Db 1269 CATGGGAGCATTTGCGCGCTGTGCGCGCGGTGATC--AGTGTGCTTACAGAGCGCA 1325  
 QY 381 nAspGlyAspPheTyrPhePheLysGlyArgCysTrpArgPheArgGlyProLysPr 401  
 Db 1326 AGACGTCGTCTTTCTTTTCAAGGTAGACCGTATGCTCTTTCGAGAAAGCAACT 1385  
 QY 401 ovalTrpGlyLeuProGln--LeuCybArgAlaGly--GlyLeuPro--ArgHisPr 418  
 Db 1386 GAGACCGCGGTACCAAGCGCGGTGACCACTATGCGTGGGATCCCTATAGACCGCAT 1445

QY 418 oAspAlaAlaLeuPhePheProProLeuArgArgLeuIleLeuPheLysGlyAlaArgTy 438  
 Db 1446 TGACACGGCGCATCTGTGTGGAGCCCAAGAGCCACCTTCTTCTTCAAGAGACGGTA 1505  
 QY 438 rTyrValIleuAlaArgGlyLysGlnValGlnProTyrTyrProAspSerLeuGlnAs 458  
 Db 1506 CTGGCGCTTCAACGAGAGACACAGCGTGGAGACCTGGGTATCCCAAGCCCATGCT 1565  
 QY 458 pTrpGlyGlyIleProGlnGlnValSerGlyAlaLeuProArgProAspGlySer--I 477  
 Db 1566 CTGGCAGGGGATCCCTGCTCCCTTAAAGGCGCTTCTGAGCAATGACGCGCTACAC 1625  
 QY 477 elIlePhePheArgAspAspArgTyrTrpArgLeuAspGlnAlaLysLeuGlnAlaThrTh 497  
 Db 1626 CTACTTCTACAAAGGACCAACAATCTGGAATTCGACAATAGCGCTTCGAGTGAAGCC 1685  
 QY 497 rSerGlyArgTrpAlaThrGlnLeuProTrpMetGlyCysTrp--HisAlaAsnSerG 516  
 Db 1686 CGGCTACCCCAAGTCCATCTGCGGAGCTTATGCTGCCAGAGCAGCATGTGAGCCAGG 1745  
 QY 516 Y 516  
 Db 1746 C 1746

## RESULT 8

US-09-521-220-10

Sequence 10, Application US/09521220

Patent No. 639348

GENERAL INFORMATION:

APPLICANT: WILK, Horst

HINZMANN, Bernd

TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES: METALLOPROTEASES, THEIR PRODUCTION AND USE

ADDRESSER: Foley &amp; Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/521,220

FILING DATE: 08-Mar-2000

CLASSIFICATION: &lt;Unknown&gt;

21-OCT-1994

21-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/704,711

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: DE 4438838.1

FILING DATE: 21-OCT-1994

APPLICATION NUMBER: DE 4409663.1

FILING DATE: 17-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: GRANADOS, Patricia D.

REGISTRATION NUMBER: 33,683

REFERENCE/DOCKET NUMBER: 26083/124

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399

TELEX: 904136

INFORMATION FOR SEQ. ID NO. 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 3530 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear



/ ORGANISM: Homo sapiens  
US-09-171-545-7

## Alignment Scores:

Pred. No.: 2,62e-38 Length: 1524  
Score: 575.00 Matches: 167  
Percent Similarity: 48.58% Conservative: 72  
Best Local Similarity: 33.94% Mismatches: 184  
Query Match: 20.29% Indels: 70  
Gaps: 19

US-10-791-980-6 (1-520) x US-09-171-545-7 (1-1524)

```

QY 39 PheLeuGluIleuTyGlyTyrLeuSerGlnValProLyAlaProThrSerThrArg 58
DB 91 TACCTTCACATATGGGTACCTACAGAGCCTCTA-----GAAGATCTAATAC 141
QY 59 Phe-----SerAPalaIleArgAlaPheGlnTPrValSerGlnLeuPro 73
DB 142 TTCAGCCAGAGATATACCGAGGCTCTGAGAGCTTTTCAGAGAGCATGAACTTCCA 201
QY 74 ValSerGlyValLeuAPhArgAlaThrLeuArgGlnMetThrArgProArgCySGlyVal 93
DB 202 GCTCAGGTCAGCTGGATGATGCCAAGAGGCGCCGACATGAGGACCTCGTTGGGCTCA 261
QY 94 ThrAPThrAsnSerTyAlaAlaTPrAlaGluArgIleSerAPheLeuAlaArgHis 113
DB 262 GAGGATCCCTCAAC-----276
QY 114 ArgThrIleMetArgGlyValArgPheAlaLeuGlnIleValAsnIleTyrIleValGln 133
DB 277 -----CAGAAAGACCTTAATACCTGTGGTGGG---CGCTGAGAAAGAAAG 321
QY 134 HisLeuSerTyAlaGlyValAsnTPrProGlnHisLeu-ArgSerArgGlnPheGlyVal 153
DB 322 CACCTGACTTCGGCATCTGAACTGCTCCCTCCACCTTCCACCCACACAGCCGGGCA 381
QY 153 AProCyAlaProProSerSerCySGlyAlaThrSerGlnArgTPrSerSerGlyValArgPr 173
DB 382 GCCCTGGCTCAAGCTTCCAGACTGAGCA-----ANGTGG-----CTCCC 423
QY 173 OGlnProGln-----AlaProLeuThrSerGlySerProSer-----Se 186
DB 424 TTGACCTTCCAGAGGTGACGCTGGTGGCTGACATCCGCTCTCTCCATGAGCGC 483
QY 186 rIlySGlyThrThrThrMetGlyTPrAlaMetProLeuMetAlaGlnIleValProTPr 206
DB 484 CAAAGCTCGTACTGT-----CNAATACTTTTGTGAGGCTGGAGAGTTCTGGCC 534
QY 206 gThrProPheLeuProArgArgGlyGluAlaHisPheAspGlnAspGluArgTPrSerIle 226
DB 535 CATGCCGAC-ATCCAGAGCTGGGCAAGTGCATCTTCACAGAAACAGATTCTGACTGA 593
QY 226 uSerIArgArgGlyArgAsnLeuPheValValLeuAlaHisGlnIleGlyHisThrIle 246
DB 594 GGGGACCTACCTGGGGTGAACCTGGCATTTGCGAGCCATAGAGTGGCCATGTCTT 653
QY 246 uGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaProTyrTyrIleArgLeuGln 266
DB 654 GGGGCTTGGGCACTCCGATATTTCCAGGCTCATGAGCCCTGAGCTTCAAGAGGCTACCG 713
QY 266 YArgAPhAlaLeuLeuSerTPrAspAPhValLeuAlaValGlnSerLeuTyrIleYsPr 286
DB 714 GCCCACCATTAAAGCTGACCCAGATGATGGCAGGATCCAGGCTCTCTATGGCAAGAA 773
QY 286 cLeuGlyIleValAlaValGlnLeuProGlyValLeuPheThrAspPheGlnIleThr 306
DB 774 G-----AGTCAAGTGAAGAGATGAGAAAGAAAGACAGAGCTGCCACT-- 822
QY 306 pAspSerTySerProGlnIleArgArgProGlnIleGlnIleProLyTyCyHisSe 326
DB 823 -----GTGCCCCCACTGGCCACAGAACCACTGATCCATG---CAGACCTTGGAGTAG 872

```

```

QY 326 rSerPheAPhAlaIleThrValAspArgGlnGlnIleuTyrIlePheLySGlySerHis 346
DB 873 TGAATCGATGCCATATGCTGTGGGCCCCCGGAGAAACCTATGCTTTTCAAGGGGACATA 932
QY 346 sPheTPrGlyValAlaAlaAspGlyAsnValSerGlnProArgProLeu-----G1 363
DB 933 TGTGTGACTGTATCAAGATTCCAGCA-----CCGGGCCCCCTTTGTTCCAGAGTTC 980
QY 363 nGluArgTPrValGlyLeuProProAsnIleGlnAlaAlaValSerLeuAsnAspG1 383
DB 981 TGGCCTTTGGAGAGGGGCTCCCGGAAACCTGAATGCTGTCTTATCTGCTCGAACACA 1040
QY 383 YAspHeTyTrPhePheLySGlyArgCySTPrArgPheArgGlyProLySProValTr 403
DB 1041 ATGATATTACTCTTTAAAGGAGACAAAGGTGGCGCTTACTTATTTCAAGATCTTCC 1100
QY 403 pGlyLeuProGlnIleuCySArgAlaGlyGlyLeuProArgHisProAPhAlaLeuPh 423
DB 1101 TGGCTTCCCAAG-----AAGCTGAATAGAGTAAGAACCTTAACCTGATGCACTCTCTA 1154
QY 423 ePheProProLeuArgArgLeuIleuPheLySGlyAlaArgTyr-----Va 440
DB 1155 TTGGCTCTCAACCAAGAGGTTCCTTTAAGGCTCCGGGTACTGGGAGTGGAGCA 1214
QY 440 lLeuAlaArgGlyGlyLeuGlnValGluProTyrTyrProArgSerLeuGlnAsp---Tr 459
DB 1215 GCTAGCCCGAACTGACTTCAGC-----AGCTACCCCAACCAATCAAGGTTTGT 1265
QY 459 pGlyGlyIleProGlnIleValSerGlyAlaLeuProArgProAspGlySerIleIlePh 479
DB 1266 TACGGAGTGCACAAACAGCCCTCGGCTGCTTGAAGTGGCAAGATGGCGAGTACTT 1325
QY 479 ePheArgAspAspArgTyTrPArgLeuAspGlnAlaLyLeuGlnAlaThrThrSerG1 499
DB 1326 CTTCAGAGGCAAGTCTACTGGCGCTCAACAG---CAGCTTCAGTAGAAGAAAGGCTA 1382
QY 499 YArgTPrAlaThrGlnLeuProTPrMetGlyCyS 510
DB 1383 TCCCAAGAAATATTTCCCAACTGATGATGACTGT 1416

```

## RESULT 10

US-09-171-545-22  
/ Sequence 22, Application US/09171545  
/ Patent No. 6566116  
/ GENERAL INFORMATION:  
/ APPLICANT: Koji, Yoshimura  
/ APPLICANT: Yuichi, Hikiuchi  
/ APPLICANT: Atsushi, Nishimura  
/ TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION AND USE  
/ CURRENT APPLICATION NUMBER: US/09/171,545  
/ CURRENT FILING DATE: 1999-07-26  
/ NUMBER OF SEQ ID NOS: 23  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 22  
/ LENGTH: 2264  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (95) ..(1618)  
US-09-171-545-22

## Alignment Scores:

Pred. No.: 4.63e-38 Length: 2264  
Score: 575.00 Matches: 167  
Percent Similarity: 48.58% Conservative: 72  
Best Local Similarity: 33.94% Mismatches: 184  
Query Match: 20.29% Indels: 70  
Gaps: 19

US-10-791-980-6 (1-520) x US-09-171-545-22 (1-2264)

```

Qy 39 PheLeuGluValYrGlyTYrLeuGlnGluValProlyAlaProThrSerThrArg 58
Db 185 TACCTGCACAAATATGGGTACCTACAGAACCTCTA-----GAAGACTCTAATAAC 235
Qy 59 Phe-----SerAlaLeuLeuArgAlaPheGlnTrpValSerGlnLeuPro 73
Db 236 TTCAGCCAGAAATATACCCAGAGCTGTGAGAGCTTTTCAGAGAACATCTGAACCTTCCA 295
Qy 74 ValSerGlyValAlaLeuAspArgAlaThrLeuArgGlnMetThrArgProArgCysGlyVal 93
Db 296 GTCTCAGCTCAGCTGATGATGCCCAAGAGGCCCATGAGGACCTCTGTGTGGCTTA 355
Qy 94 ThrArgThrAsnSerTYrAlaAlaTrpAlaGluArgGlieSerAspLeuPheAlaArgHis 113
Db 356 GAGAGTCCCTTCACAC-----370
Qy 114 ArgThrIleMetArgArgLysLysLeuArgPheAlaLysGlnGlyAsnLysTrpTYrIleGln 133
Db 371 -----CAGAAAGACCTTAAATACCTGTGGTGGC---CGCTGAGAAAGAG 415
Qy 134 HisLeuSerTYrArgLeuValAsnTrpProGlnHisLeu-ArgSerArgGlnPheGlyAla 153
Db 416 CACCTGACTTTCCGATCTTGAACCTGCCCTCCACCTTCCACCCACACAGCCGGGCA 475
Qy 153 AProCysAlaProProSerSerCysGlyAlaThrSerGlnArgTrpSerSerGlyArgPr 173
Db 476 GCCCTGCCTCAAGCCTTCCAGAGCTGAGCA-----ATGTGG-----CTCCC 517
Qy 173 GlnProGln-----AlaProLeuThrSerGlySerProSer-----Se 186
Db 518 TTGACCTTCCAAAGAGCTGAGCTGGCTGGCTGACATCCGCTCTCTCCATGAGCCGC 577
Qy 186 rLyLeGlyThrThrThrMetGlyTrpAlaMetProLeuMetAlaGlnGlyAlaProTrpAr 206
Db 578 CAAGACTGTACTGT-----CCAATCTTTTATGGGCCCTGGAGAGTCTGGCC 628
Qy 206 gThrProPheLeuProArgArgGlyGluAlaHisPheArgGlnAspGluArgTrpSerLe 226
Db 629 CATGCCGAC-ATCCAGAGCTGGGAGAGTGCATTTCCAGCAAGACGAGTTCTGGACTGA 687
Qy 226 uSerArgArgArgGlyArgAsnLeuPheValValLeuAlaHisGluIleGlyHisLeu 246
Db 688 GGGAGCTTACCGTGGGGTGAACCTGGCATATTCACACCCATGAAGTGGCCAGCTCT 747
Qy 246 uGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaProTYrTYrArgLeuGlu 266
Db 748 GGGGCTTGGGACCTCCGATATTCACAGCCTTCAAGCCCAAGCTTCAAGAGGCTACCG 807
Qy 266 yArgAspAlaLeuLeuSerTrpAspArgValLeuAlaGlnSerLeuTYrGlyLysPr 286
Db 808 GCCCCTTTAAGCTGCACCCAGATGATGTGGCAGAGTCCAGGCTCTCTATGGCAAGA 867
Qy 286 cLeuGlyGlySerValAlaValGlnLeuProGlyLysLeuPheThrAspPheGluThrTr 306
Db 868 G-----AGTCCAGTGAATAAGGATGAGAAAGAAAGAGAGAGAGCTGCCACPr- 916
Qy 306 pAspSerTrpSerProGlnGlyArgArgProGluThrGlnGlyProLYrCysHisSe 326
Db 917 -----GTGCCCCCAAGTCCCAAGAACCCAGTCCCATG---CCAGACCTTGCAGAG 966
Qy 326 rSerPheAspAlaIleThrValAspArgGlnGlnLeuTYrIlePheLysGlySerHis 346
Db 967 TGAAGTGAATGCAGATGATGCGGAGCCGATGGGAAAGCTATGCTTTCAAGGGGAGCTA 1026
Qy 346 sPheTrpGluValAlaAlaAspGlyAsnValSerIleProArgProLeu-----GI 363
Db 1027 TGTGTGACTGTATCAGATTCAGAG-----CCGGGACCTTGTTCGAGTGC 1074
Qy 363 nGluArgTrpValGlyLeuProProAsnIleGluAlaAlaValSerLeuAsnAspGlu 383
Db 1075 TGCCCTTTGGAGAGGGGCTCCCGGAAACCTGAGATCTGTCTACTGCGCTTCAAGACA 1134
Qy 383 yAspPheTrpPhePheLysGlyArgCysTrpArgPheArgGlyProLYrProValTr 403

```

```

Db 1135 ATGATTCACCTCTTTAAGGAGACAAGGTGTGGCTTACATTAATTCAAGATGTCTCC 1194
Qy 403 pGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHisAspAspAlaAlaLeuPh 423
Db 1195 TGCGTTCCTCCAG-----AGCTGAATAGGTGTGAACCTTAACCTGATGACGCTCTTA 1248
Qy 423 ePheProProLeuArgArgLeuIleLeuPheLysGlyAlaArgTYr-----Va 440
Db 1249 TTGGCTCTTCAACCAAAAGGTGTCTTTTAAGGCTCCGGGTACTGAGTGGAGCA 1308
Qy 440 lLeuAlaArgGlyGlyLeuGlnValGluProTYrTYrProArgSerLeuGlnAsp---Tr 459
Db 1309 GCTAGCCCGAAGCTGACTTCAC-----AGCTTACCCCAACCAATCAAGGGGTGTGT 1359
Qy 459 pGlyLYrLeuProGlnGluValSerGlyAlaLeuProArgProAspGlySerIleIlePh 479
Db 1360 TACGGAGTGCACAAACAGCCTCGCTGTATAGTTGGCAAAATGGCCAGATCTACTT 1419
Qy 479 ePheArgAspAspArgTYrTrpArgLeuAspGlnAlaLysLeuGlnAlaThrThrSerGI 499
Db 1420 CTTCAAGGGCAAAGTCTACTGGCGCTCAACAG---CAGCTTCAAGTAAAGAAAGCTA 1476
Qy 499 yArgTrpAlaThrGluLeuProTrpMetGlyCys 510
Db 1477 TCCCAAAATATTTCCCAACATGATGCACTGT 1510

RESULT 11
US-09-171-545-23
; Sequence 23, Application US/09171545
; Patent No. 6566116
; GENERAL INFORMATION:
; APPLICANT: Koji, Yoshimura
; APPLICANT: Yuichi, Yoshichi
; APPLICANT: Atsushi, Nishimura
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION AND USE
; FILE REFERENCE: 48712/342
; CURRENT APPLICATION NUMBER: US/09/171, 545
; CURRENT FILING DATE: 1999-07-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (90)..(1640)
; US-09-171-545-23

Alignment Scores:
Pred. No.: 7,08e-37 Length: 2049
Score: 560.00 Matches: 173
Percent Similarity: 44.97% Conservative: 73
Best Local Similarity: 31.63% Mismatches: 194
Query Match: 19.76% Indels: 108
DB: 4 Gaps: 18

US-10-791-980-6 (1-520) x US-09-171-545-23 (1-2049)
Qy 8 LeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHisLeuAspAlaGln-----23
Db 54 CTCAGGAGAGCTTGGAACTCTGTCTGTGGTGGACATGACTGGACAGCTGTGGCTG 113
Qy 24 -----ProAlaGluArgGlyGlyGlnGlnLeu-----ArgLysGluAla 36
Db 114 GCCTTCTACTTCTGTGACAGTCTCAGGCGGCTCTGGGCGCTGCAAGAGAGAGCG 173
Qy 37 GluAla---PheLeuGluLysTYrGlyTYrLeuGlnGlnValProLYrAlaProThr 55
Db 174 GTGGATTAATCCTGTTCAGATATGAGTATTCAGAAACCTCTCGAAGAGCT---GAT 230
Qy 56 SerThrArgPheSerAsp-----AlaIleArgAlaPheGlnTrpValSerGlnLeu 72

```

```

Db      231 GACTTCAGGCTGAGAAATATCAAGAGGCTCTAAGAACTTCCAGGAAGATCTGAAGTGC 290
Qy      73 ProValSerGlyValLeuAspArgAlaThrLeuArgGlnMetThrArgProArgCysGly 92
Db      291 CCTGTTCCGGTCATATGATGATGATCCCAAGAGGCCGCTATGAAAGCAGCCCGCTGTGGCC 350
Qy      93 ValThrAspThrAsnSerTyrAlaAlaTrrAlaGluArgLLeuSerAspLeuPheAlaArg 112
Db      351 CTGGAGGATCCTTTCAC- 368
Qy      113 HisArgThrLysMetArgArgLysLysAspPheAlaLysGlnGlyAsnLysTrrTyrLys 132
Db      369 -----CAGAACTCTGAAATATCTGCTTTCGGCCAC--TGAGAAAG 410
Qy      133 GlnHisLeuSerTyrArgLeuValAsnTrrProGlnHisLeuArgSerArgGlnPheGly 152
Db      411 AAGCACTTGACATCCCATCTTGACATGACGTCGCCCTCCACCTCTCACCCCTCCAGAGTCCGA 470
Qy      153 Ala-ProCysAlaAlaProProSerSerCysGlyAlaAlaThrSerGlnArgTrrSerSerGlyAr 172
Db      471 GCACCCCTGCACTACAGCCTTTAACTATTTGAGCAATGTAAGCCCTGACCTTCCGGAG 530
Qy      172 gProGlnProGlnAlaProLeuThr--SerGlySerProSerSerGlyThrThr 191
Db      531 GTGAAAGCTGGTGGGCTGATATCCGCTCTGCTTCATGAGCCGCAAGCCCATATCTGC 590
Qy      191 rMetGlyTrrAlaMetCProLeuMetAlaGlnGlyAlaProTrrArgThrProPheLeuPr 211
Db      591 T-----CNAACGCTTGTATGGGCTCGGAGAGGAGCTCTGCC- CAGTCGAGAGTCC 640
Qy      211 oArGArGlyGlyAlaHisPheAspGlnAspGluArgTrrPrrSerLeuSerArgArgGly 231
Db      641 AGAGCTTGGAGAGTACACTTGATTAACATGATTTCTGACCGAGGCACTTACCAAGG 700
Qy      231 YArgAsnLeuPheValValLeuAlaHisGlnLysLeuGlyHisThrLeuHisGly 251
Db      701 AGTGAACCTAACGATCATTTGCGGCCCATAGAGTGGCCACGCGCTGGGACCTTGGGCAATTC 760
Qy      251 rProAlaProArgAlaLeuMetAlaProTrrTrrLysArgLeuGlyArgAspAlaLeuLe 271
Db      761 CGGATATACCACGAGCACTATGCGCTGTTTACGCTGCTACCAAGCCCTTACTTCAGGCT 820
Qy      271 uSerTrrAspAspValLeuAlaValGlnSerLeuTrrGlyLysProLeuGlyGlySerVa 291
Db      821 GCATCCGATGATGTGGCAGGATCCAGCGCTCATGCAAG- 863
Qy      291 lAlaValGlnLeuProGlyLysLeuPheThrAspPheGlnThrTrrAspSerTrrSerPr 311
Db      863 ----- 863
Qy      311 oGlnGlyArgArgProGlnThrGlnGly- 320
Db      864 -----AGAGAGCGCGGAGCCAGAAAGAGAGAGAGAGTGGAGATGCACCTGTCTC 320
Qy      321 -----ProLysTrrCysHisSerSerPheAspAlaAla 331
Db      917 AACAGTGCACAAACCAGTCCATGCCAAACCTGCGAGCAGTGAATGATGCCAT 976
Qy      331 eThrValAspArgGlnGlnGlnLeuTrrLysPheLysGlySerHisPheTrrGluValAl 351
Db      977 GATGCTAGGGGCTCGGGGAAAGACCTATGCTTTAAGGATGATGCTGCTGAAC 1036
Qy      351 aAlaAspGlyAsnValSerGlnProArgProLeu-----GlnGluArgTrrValG 368
Db      1037 AGATTCAAGG-----CCAGGCGCTTGTTCGAGTCTGCGCTTTGGAGGG 1084
Qy      368 YLeuProProAsnLLeGlnAlaAlaValSerLeuAspAspGlyAspPheTrrPhePh 388
Db      1085 GCTTCTCGAAACCTGAGTCTGTCTACTCTCCCGGACACAGCGAGCATTTCTT 1144
Qy      388 eLysGlyGlyArgCysTrrArgPheArgGlyProLysProValTrrGlyLeuProGlnLe 408

```

```

Db      1145 CAAGGAAACAAGTGTGGCGGTATGTGATTTCAAGTTGTCTCGCTTCCCATG-- 1202
Qy      408 uCysArgAlaGlyGlyLeuProArgHisProAspAlaAlaLeuPhePheProLeuAr 428
Db      1203 -----AAACTCAACAAGAGTGAACCCAACTTGAATGACGTCTTAACTGTTAACT 1258
Qy      428 gArgLeuLLeuPheLysGlyAlaArgTrrTrr-----ValLeuAlaArgGlyG 445
Db      1259 GAAGGTGTTCTTTTAAAGGCTCAGATCTGGCAATGGAGTGAATGACTGACCAACTGA 1318
Qy      445 YLeuGlnValGluProTrrTrrProArgSerLeuGlnAsp---TrrGlyGlyLeuProG 464
Db      1319 CCTAGT-----CGCTTACCCCAACCAATCAAGAACTTTCACTGAGAGTCCAGA 1369
Qy      464 uGluValSerGlyAlaLeuProArgProAspGlySerLLeuPhePheArgAspAspAr 484
Db      1370 CCAACCTCAGCAGCACTTGAAGCTGGCAAGATGGCAAGTCTACTTCTTCAAGGCAAAAG 1429
Qy      484 gTrrTrrArgLeuAspGlnAlaLysLeuGlnAlaThrThrSerGlyArgTrrAlaThrG 504
Db      1430 GTACTGGCGCTTAAACGACAACTTGCAGTGGCAAGGCTATCCAGAAATGACACACA 1489
Qy      504 uLeuProTrrMetGlyCys 510
Db      1490 C-----TGATGCACTGT 1502

RESULT 12
US-08-704-711A-9
Sequence 9, Application US/08704711A
Patent No. 6114159
GENERAL INFORMATION:
APPLICANT: WILLY, Horst
APPLICANT: HINZMANN, Bernd
TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
NUMBER OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
CORRESPONDENCE ADDRESSES:
ADDRESSER: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/704,711A
FILING DATE: 20-NOV-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: WO PCT/DE95/00357
FILING DATE: 17-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE 4438838.1
FILING DATE: 21-OCT-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE 4409663.1
FILING DATE: 17-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 26093/124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3437 base pairs
TYPE: nucleic acid

```

```

/ STRANDEDNESS: single
/ TOPOLOGY: linear
us-08-704-711A-9

Alignment Scores:
Pred. No.: 3,896-36 Length: 3437
Score: 555.00 Matches: 161
Percent Similarity: 45.60% Conservative: 67
Best Local Similarity: 32.20% Mismatches: 220
Query Match: 19.58% Indels: 53
DB: 3 Gaps: 17

US-10-791-980-6 (1-520) x US-08-704-711A-9 (1-3437)

QY 37 GIUAlaPheLeuGluYerTYr-----LeuAnaGlunValPro 51
DB 219 GAGCCTGCTACAGCAATATGCTACCTGCCCGGGAGACTCGTACCAACACACAG 278
QY 52 LybAlaProThSerThArPheSerAaIleArgAlaPheGlnTPValSerGln 71
DB 279 CGCTCACCCAGTCA-----CTCTCAGGGCCATCGCTGCATGACAGAGTTTACGAG 332
QY 72 LeuProValSerGlyValLeuAaPaRgAlaThrLeuAaGlnMerThArGProArG 91
DB 333 TTGCAAGTAAACAGGAAAGCTGATGACACACATGAAAGCCATGAGGCGCCGATGT 392
QY 92 GIYAlaIThrAaPThArAaSerTyraIaIaITrPaIaGluArgIleSerAaPheAla 111
DB 393 GGTGTTCCAGACAG-----TTGGG 413
QY 112 ArgHlaaGThLyMeArGArgLybAaRgPheAlaYbGlnGlyAaMlyeTPTy 131
DB 414 GCTGATCAAGAGCCAAATGTTGAAAGAAAGGCTACGCGATCAGAGGTCTCAATGGCA 473
QY 132 LybGlnHlaeLeuSerTyraRgLeuValAaNTTP-ProGlnHlaeAaRgSerArgGln 151
DB 474 CATAAATGAATCATCTTCTGATCCAGATTAACACCCCAAGGTGGCGAGTATGCCA 533
QY 151 eGlyAlaProCybAlaProProSerSerCybGlyAlaThrSerGlnArgTPSerSerG 171
DB 534 TACAGAGCCATTGCGAAGCGCTTCCGCGTGGAGAGTGCACACACACTGCGCTTCCG 593
QY 171 YArg---ProGlnProGlnAlaProLeuThrSerGlySerPro-----SerSe 186
DB 594 GAGGTGCCCTTATGCTACATCGGTGAGGCGCATGAAAGCAGCGACATCATGATCTTC 653
QY 186 rLybGlyThrThrMetGlyTrPaIaMePProLeuMeAlaGlnGlyAlaProTPAr 206
DB 654 TTTGCCAGAGGCTTCCATGCGACAGACAGCCCTTCGATGAGTGGAGGCGGCTTCCG 713
QY 206 gThrProPheLeuProArG-----ArgGlyGluAlaHlaPheAaRgAlaPgluAr 223
DB 714 CATGCCCTAC-TTCCAGGCGCCCAACATGAGAGAGACCCACTTGTGATCTGCCAGGC 772
QY 223 gTPSerLeuSerArGArg-----ArgGlyArgAaLeuPheValaIleuAlaHla 241
DB 773 TTGACATGTCAGAGATAGAGATCTGAATGAATGACATCTTCTGTGCTGCTGACAGA 832
QY 241 uIlleGlyHlaeThrLeuGlyLeuThrHlaSerProAlaProArGAlaLeuMeAlaProTy 261
DB 833 GCTGGGCGCATGCTCGGGGCTCGACATTCAGATCCCTCGGCTCATGAGCACCTT 892
QY 261 rTyTybArGLeuGly---ArgAaPaIaLeuLeuSerTPAaPaPaValaIleuAlaVal 280
DB 893 TTACCAAGTGAAGACAGAGAAATTTGTGCTGCCCATGATGACCCGCGGAGCATCA 952
QY 280 nSerLeuTyrgLybProLeuGlyGlySerValaIaValaGlnLeuProGlyLyLeuPh 300
DB 953 GCAACTTATGGGGGAGTCAAGG-----TTCCCAACAAGATGCC 994
QY 300 eThrAaPheGluThrTPAaPSeTySer---ProGlnGlyAaRgArGProGlnThrG 319
DB 995 CCTCAACCCAGACTACTCCGGCGCTTCTGTCTGTATAAACCAAAACCCACACTTA 1054

```

```

QY 319 nGlyProLySerTyrcYbHlaSerSerPheAaPaIaIleThrValaAaRgGlnGlnLe 339
DB 1055 TGGGCCAAACATCTGTACGGGAACCTTGAACCGGTGCATGCTCCGA---GGGAGAT 1111
QY 339 uTyTrIlePheLybGlySerHlaPheTPGluVal-----AlaIaAaPglyAa 355
DB 1112 GTTGTCTTCAAGAGCGCGCTGTTGCGGGGTGAATTAACCAAGATGATGATGA-- 1169
QY 355 nValSerGluProArGProLeuGlnGluArgTrPaIaGlyLeuProProAmIleGluAl 375
DB 1170 -----TACCAATGCCATTTGGCCTTCTGCGGGGCGCTTCTGCTCATAC-- 1220
QY 375 AlaIaIaValSerLeuAaAaPglyAaPheTyPhePheLybGlyArgCybTPAr 395
DB 1221 -ACGCTTAAGAGAGAAAGATGGAATGGAATTTGTTCTTCAAGAGACAGATTTGGT 1279
QY 395 gPheArgGlyProLybProValTPrglyLeuProGlnLeuCybArg-----AlaGly 413
DB 1280 GTTTGATGAGCGCTCCCTGGAACCTGCTACCCCAAGACATTAAGAGCTGGCGGAG 1339
QY 413 YLeuProArGHisPro---AaPaIaIaLeuPhePheProProLeuAaRgArGLeuIle 432
DB 1340 GCTGCTACCGACAGATGATGCTGCTCTCTTCTGATGCCCAATGAGAAAGACTTACT 1399
QY 432 uPheLybGlyAlaArgTyTyTyValaIleuAlaArgGlyGlyLeuGlnValaGluProTy 452
DB 1400 CTTCCTGGAACAAAGTACTACCTTCCGTTCAAGAGAGCTCAGGCGGAGTATGAGATGA 1459
QY 452 rProArGSerLeuGlnAaPTrPgLylyIleProGlnGluValSerGlyAlaLeuProAr 472
DB 1460 CCCCAAGAACTCAAAAGCTGTGGAAAGGATCTGTAGTCTCCAGAGGCTCATTCATGG 1519
QY 472 gProAaPglySerIle---IlePhePheArgAaPaRgTyTrTPArGLeuAaPgluAl 491
DB 1500 CAGCATGAAGCTTCACTTCTTCTTACAGGGGAGCAAAATCTGAAATTCACACACA 1579
QY 491 aLybLeuGlnAlaThrThrSerGlyArgTrPaIaThrGluLeuProTPMeGlyCy 510
DB 1580 GAGCTGAAGGTAAAGACCGGGCTACCCCAAGTCAAGCCTGAGGAGCTGATGGCTGC 1637

RESULT 13
US-09-521-220-9
/ Sequence 9, Application US/09521220
/ Patent No. 639348
/ GENERAL INFORMATION:
/ APPLICANT: WILL, Horst
/ HINZMANN, Bernd
/ TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
/ METALLOPROTEASES, THEIR PRODUCTION AND USE
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Foley & Lardner
/ STREET: 3000 K Street, N.W., Suite 500
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20007-5109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/521,220
/ FILING DATE: 08-Mar-2000
/ CLASSIFICATION: <Unknown>
/ 21-OCT-1994
/ 17-MAR-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/704,711
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: DE 4438838.1

```

```
/
/ FILING DATE: 21-OCT-1994
/ APPLICATION NUMBER: DE 4409663.1
/ FILING DATE: 17-MAR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: GRANADOS, Patricia D.
/ REGISTRATION NUMBER: 33,683
/ REFERENCE/DOCKET NUMBER: 26083/124
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 672-5300
/ TELEFAX: (202) 672-5399
/
/ TELEEX: 904136
/
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3437 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 9:
/
US-09-521-220-9
/
Alignment Scores:
Pred. No.: 3,89e-36 Length: 3437
Score: 555.00 Matches: 161
Percent Similarity: 45.60% Conservative: 67
Best Local Similarity: 32.20% Mismatches: 220
Query Match: 19.58% Indels: 53
DB: 3 Gaps: 17
/
US-10-791-980-6 (1-520) x US-09-521-220-9 (1-3437)
/
QY 37 GUAUAApHeuEngluuYrtyrGlyTyr-----LeuansgluInuValPro 51
Db 219 GAAGCCTGGCTACAGCAATATGGCTTACCTGCCCTCCGGGAGACTTACCAACACACAG 278
QY 52 UvAAlaProThSerThraRphSerAbaAlaileArgAlaPheGlnTrpValSerGln 71
Db 279 CCGTCAACCCCAAGTCA-----CTCTCAGCGCGCATCGCTGCATGCACAGATTTCAGGC 332
QY 72 LeuProValSerGlyValLeuAbaRphAlaThrluAaArglInuMetThraRphProArgCys 91
Db 333 TTGCAGATTAACAGCAAGCAAGCTATGCAACACCAATGAGGCCATGAGCGCCCGCATGT 392
QY 92 GUAUAApHeuEngluuYrtyrGlyTyr-----LeuansgluInuValPro 51
Db 393 GGTGTTCAGACAG-----TTTGGG 413
QY 112 ArgHlaBArgThrluMetArgAlaGlyValArgPheAlaGlyValAsnLysTrpTyr 131
Db 414 GCTGAGATCAAGGCCAATGTTGGAAGAGCGCTACGCCATCCAGGGTCTCAAAATGGCA 473
QY 132 UvAAlaProThSerThraRphSerAbaAlaileArgAlaPheGlnTrpValSerGln 71
Db 474 CATATGAAATCACTTTTGCATCCAGAAATTAACACCCCAAGTGGCGAGATATGCCA 533
QY 151 eGlyAlaProCysAlaProProSerSerCysGlyAlaThrSerGlnArgTrpSerSerGln 171
Db 534 TACGAGGCGCATTCGCAAGCGCTTCGCGGTGTGGAGAGTGCACACCACTGCCCTTCCG 593
QY 171 YArg---ProGlnProGlnAlaProLeuThrSerGlySerPro-----SerSe 186
Db 594 GAGGTGCCCTTATGCTTACATCCGTGAGGCCATGAGAGCGGCCGACATCATGATCTTC 653
QY 186 rLysGlyThrThrluMetGlyTrpAlaMetProLeuMetArglInuValProTrpPar 206
Db 654 TTGGCCGAGGCTTCATGCGAGACAGACGCCCTTCATGAGTGTGAGGCGGCTTCCGGCC 713
QY 206 gThrProPheLeuProArg-----ArgGlyGluAlaHisPheAspGlnAspGluArg 223
Db 714 CATGCCATAC-TTCCAGGCGCCCAATGAGAGAGACCACTTTGACCTTCCGAGGCC 772
QY 223 gTrpSerLeuSerArgArg-----ArgGlyArgAsnLeuPheValValLeuAlaHisGln 241
Db 773 TTGGACTGTCAAGAAATGAGATCTGAATGAAATGACATCTTCTGTGTGCTGTGACAGA 832
```

```
QY 241 uLLeGlyHISThrLeuEnglyLeuThrluSerProAlaProArgAlaLeuMetAlaProTy 261
Db 833 GCTGGCCCATGCGCTTGGGGCTTCAGACATTCACAGTACCCCTGCGGCATCATGAGCACCTT 892
QY 261 rTyrluAbaRphLeuGly---ArgAbaAlaLeuLeuSerTrpAbaRphValAlaValGln 280
Db 893 TTACAGAGGATGACACAGCAAGAAATTTGTCTCTGCGCATATGACCGCGGCGCATCA 952
QY 280 nSerLeuTyrluYrtyrProLeuEnglySerValAlaValGlnLeuProGlyLysLeuPh 300
Db 953 GCACCTTATGAGGGGTGAGTCAGG-----TTCCCAACCAAGATGCC 994
QY 300 eThraRphHeGluThrluThraRphSerTrpSer---ProGlnLysArgArgProGluThrgl 319
Db 995 CCGTCAACCCAGACATCACTCCGCGCTTCTGTTCTTGATTAACCCCAAAACCCACCTTA 1054
QY 319 nGlyProLysTyrluYrtyrSerHisSerSerPheAbaAlaileThrValAbaRphGlnGlnLe 339
Db 1055 TGCGCCCAACATCTGTACCGGAACTTGAACCGGTGCCATGCTCCGA---GGGAGAT 1111
QY 339 uTyrluPheLysGlySerHisPheTrpGluVal-----AlaAlaAspGlyAs 355
Db 1112 GTTGTCTTCAAGAGCGCTGTCTGCGCGGTGAGGAATTAACCAAGTATGATGGA-- 1169
QY 355 nValSerGluProArgProLeuEngluArgTrpValGlyLeuProProAsnLleGluAl 375
Db 1170 -----TACCAATGCCCATTTGGCCAGTTCTGCGGGGCTCTGCTGCTCAATCAAC-- 1220
QY 375 aAlaAlaValSerLeuAbaRphGlyAbaRphTyrluPhePheLysGlyGlyArgCysTrpPar 395
Db 1221 -ACTGCCCTACAGAGAGAGAGATGGCAAAATTCGTTCTTCAAGAGAGACAGATTTGGGT 1279
QY 395 gPheArgGlyProLysProValTrpGlyLeuProGlnLysCysArg-----AlaGlyGln 413
Db 1280 GTTGTATGAGGCGCTCCCTGGAACCTGCTACCCCAAGCAATTAAGAGCTGGCGCAGAG 1339
QY 413 ULeuProArgHisPro---AspAlaAlaLeuPhePheProProLeuAbaRphLeuLle 432
Db 1340 GCTGCCATACGCAACAAATGATGATCTCTCTCTTGTGATGCCCAATGAAAGACCTACTT 1399
QY 432 uPheLysGlyAlaArgTyrluYrtyrValLeuAlaArgGlyGlyLeuGlnValGluProTyrlu 452
Db 1400 CTTCCGTGAAACCAATGACTTACCGTTTCAACAGAGAGCTCAGGGCAGTGGATAGCGAGTA 1459
QY 452 rProArgSerLeuGlnAbaRphTrpGlyGlyLleProGluGluValSerGlyAlaLeuProAr 472
Db 1460 CCCCAGAAACATCAAAAGTCTGGGAAGGATCCCTGAGTCTCCAGAGGGTCAATTCATGGG 1519
QY 472 gProAbaRphLysSerIle---IlePhePheAbaRphAbaRphTyrluTrpArgLeuAbaRphAla 491
Db 1520 CAGCGATGAAGTCTTCACTTACTTCAACAGGGGAAACAAATCTGAAATTCACACACA 1579
QY 491 aLysLeuGlnAlaThrThrluSerGlyArgTrpAlaThrGluLeuProTrpMetGlyCys 510
Db 1580 GAAGCTGAAGGTGAAGACCGGGCTACCCCAAGTCAAGCCCTGAGGAGCTGATGGCTGC 1637
/
RESULT 14
/ US-09-919-497-34
/ Sequence 34, Application US/09919497
/ Patent No. 6773883
/
/ GENERAL INFORMATION:
/ APPLICANT: Multer, George L.
/ TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
/ FILE REFERENCE: B0801/7225
/ CURRENT APPLICATION NUMBER: US/09/919,497
/ CURRENT FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: US 60/221,735
/ PRIOR FILING DATE: 2000-07-31
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 34
/ LENGTH: 3437
```



```

TYPE: DNA
ORGANISM: Homo sapiens
US-09-919-497-34

Alignment Scores:
Pred. No.: 3,896-36 Length: 3437
Score: 555.00 Matches: 161
Percent Similarity: 45.60% Conservative: 67
Best Local Similarity: 32.20% Mismatches: 220
Query Match: 19.58% Indels: 53
DB: 4 Gaps: 17

US-10-791-980-6 (1-520) x US-09-919-497-34 (1-3437)

QY 37 GluAlaPheLeuGluLeuYrGlyYr-----LeuAenGluGlnValPro 51
DB 219 GAAGCTGGCTACACATATGCTACCTGCTCCCGGGAGCTACGATCCACACACAG 278
QY 52 LybAlaProThrSerThrArgPheSerApaAlaIleArgAlaPheGlnTrpValSerGln 71
DB 279 CGCTCAGCCCAAGTCA-----CTCTCAGCGGCATCGCTGCATGCAGAAATTTCAGGC 332
QY 72 LeuProValSerGlyValLeuAaPargAlaThrLeuArgGlnMetThrArgProArgCys 91
DB 333 TTGCAAGTAACAGGCAAGCTGTATGACACACCATGAGGCCATGAGGCGCCCGATGT 392
QY 92 GlyValThrAaPThrAaSerSerTyraIaIaIaTrpAlaGluArgIleSerAaPheAla 111
DB 393 GGTGTTCCAGACAG-----TTTGGG 413
QY 112 ArgHlaArgThrLyMetArgArgLybArgbArgPheAlaLyGlnGlyAaAlaYrTrpYr 131
DB 414 GCTGATATCAAGGCAATGTTGAAAGAAAGCGTACGCTACAGGATCAAGATGCA 473
QY 132 LybGlnHlaSerSerTyraArgLeuValAaTrpProGlnHlaSerAaArgGlnPhe 151
DB 474 CATAAATGAATCATCTTGTGATCCAGAAATTAACCCCGCAAGGTGGCGAGTATGCCA 533
QY 151 eGlyAlaProCysAlaProProSerSerCysGlyAlaThrSerGlnAaGTrpSerSerG 171
DB 534 TACGAGCGCATTCGCAAGCGCTTCGCGGTGGAGAGTCCACACACATCGCGCTTCCC 593
QY 171 YArg---ProGlnProGlnAlaProLeuThrSerGlySerPro-----SerSe 186
DB 594 GAGGTGCGCTATGCTCATCGTGAAGGCGCATGAGAGCAGCGGCATCATGATCTTC 653
QY 186 PLYbGlyThrThrThrMetGlyTrpAlaMetProLeuMetAlaGlnGlyAlaProTrpAr 206
DB 654 TTTGCCAGAGGCTTCATGCGACAGCAGCGCTTCGATGTGAGGGCGGCTTCCGTGCC 713
QY 206 GThrProPheLeuProArg-----ArgGlyGlnAlaHlaSphaAaPglAaPglAa 223
DB 714 CATGCTTAC-TTCCAGGCGCCCAACATTTGAGAGAGACCCACTTTGACTTCGCGAGCC 772
QY 223 gTrpSerLeuSerAaArg-----ArgGlyArgAaPheAaPheValAlaLeuAlaHlaS 241
DB 773 TTGAGCTGACAGAAATGAGATGTGAATGAATGACATCTTCGTGTGCGCTGACAGA 832
QY 241 uIlleGlyHlaThrLeuGlyLeuThrHlaSerProAlaProAlaAlaLeuMetAlaProTy 261
DB 833 GCTGGGCGCATGCTCGGGGCTCGAGCATTCAGAGACCCCTCGCATCATGAGCCCTT 892
QY 261 rTyTrpAaPLeuGlyLy---ArgAaPAlaLeuLeuSerTrpAaPValAlaLeuAlaVal 280
DB 893 TTACCAAGTGAATGACAGAGAAATTTTGCTGCGCATGATGACCGCGGGGATCCA 952
QY 280 nSerLeuTyrgLybProLeuGlyLySerValAlaValGlnLeuProGlyLyLeuAaP 300
DB 953 GCACCTTATGGGGTGAAGTCAAGG-----TTCCCGCACAAAGATGCC 994
QY 300 eThrAaPheGluThrTrpAaPSerTySer---ProGlnGlyArgArgProGlnThrG 319
DB 995 CCTCAACCCAGAGACTACCTCCGGCGCTTGTCTGTATAAACCAAAACCCACCTTA 1054

```

```

QY 319 nGlyProLybTyrgCysHlaSerSerPheAaPAlaIleThrValAaPArgGlnGlnLe 339
DB 1055 TGGGCCCAACATCTGTACGGGAACCTTGACACCGGTGCGCATGTCCGA---GGGAGAT 1111
QY 339 uTyTrlPheLeuGlySerHlaSphaTrpGluVal-----AlaAlaAaPglAa 355
DB 1112 GTTGTCTTCAAGGACCGCTGTGCGGGGTGAGAAATTAACCAAGTATGATGAGA-- 1169
QY 355 nValSerGluProArgProLeuGlnGlnArgTrpValGlyLeuProProAaHlaGlnAl 375
DB 1170 -----TACCAATGCCCATTTGGCACTTGTGGGGGGCTTGGCTGCTCATAC-- 1220
QY 375 aAlaAlaValSerLeuAaAaPglAaAaPheTyrPhePheLeuGlyLyArgCysTrpAr 395
DB 1221 -ACGCTTAACAGAGAGAAATGGAATTTGCTCTTCAAGAGACAGCATTTGGGT 1279
QY 395 gPheArgGlyProLybProValTrpGlyLeuProGlnLeuCybArg-----AlaGlyG 413
DB 1280 GTTGTATGAGGCGTCCCTGGAACCTGCTACCCCAAGACATTAAGAGCTGGCGCAGG 1339
QY 413 YLeuProArgHlaSPro---AaPAlaAlaLeuPhePheProProLeuAaArgAlaLeu 432
DB 1340 GCTGCTTCCGACAAATGATGATGCTGCTCTTCTTGATGCCAATGAGAAACCTACTT 1399
QY 432 uPheLybGlyAlaArgTyTrpValLeuAlaArgGlyLyLeuGlnValGluProTyTr 452
DB 1400 CTTCGTGGAACAAAGTACTACCGTTTCAGAGAGCTCAGGCGGAGTATGAGAGTA 1459
QY 452 rProArgSerLeuGlnAaPTrpGlyLyIleProGlnGlnValSerGlyAlaLeuProAr 472
DB 1460 CCCCAAGAAACATCAAAAGTCTGGAGAGGATCCCTGAAGTCTCCAGAGGATCATTCATGG 1519
QY 472 gProAaPglYserIle---IlePhePheArgAaPargTyTrpArgLeuAaPglAla 491
DB 1520 CAGCATGAAGTCTTCACTTACTTCTTACAGAGGAAACAAATCTGAAATTCACACCA 1579
QY 491 aLybLeuGlnAlaThrThrSerGlyArgTrpAlaThrGluLeuProTrpMetGlyCys 510
DB 1580 GAAGCTGAAGTAAGTAAGAGGCGTACCCAGATCAGCGCTGAGGAGCATGTGGCTGC 1637

RESULT 15
US-08-448-489-2
Sequence 2, Application US/08448489
Patent No. 6184022
GENERAL INFORMATION:
APPLICANT: SEIKI, Motoharu
APPLICANT: SATO, Hiroshi
APPLICANT: SHINGAWA, Akira
TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
FILE REFERENCE: 55-290P
CURRENT APPLICATION NUMBER: US/08/448,489
CURRENT FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 3403
TYPE: DNA
ORGANISM: Homo sapiens
US-08-448-489-2

Alignment Scores:
Pred. No.: 4,646-36 Length: 3403
Score: 554.00 Matches: 161
Percent Similarity: 45.60% Conservative: 67
Best Local Similarity: 32.20% Mismatches: 220
Query Match: 19.55% Indels: 53
DB: 3 Gaps: 17

US-10-791-980-6 (1-520) x US-08-448-489-2 (1-3403)

QY 37 GluAlaPheLeuGluLeuYrGlyYr-----LeuAenGluGlnValPro 51

```

Db 217 GAAGCTGGCTACAGCAATATGGCTACCTCCCGGGACCTACGTACCAACACAG 276  
 QY 52 LysAlaProThrSerThrsPheSerAerAlaIleArgAlaPheGlnTrpValSerGln 71  
 Db 277 CCTCACCCCACTCA-----CTCTCAGCGCCATCGCTCCATGCAACAAGATTTCAGCG 330  
 QY 72 LeuProValSerGlyValLeuAspArgAlaThrLeuArgGlnMetThrArgProArgCys 91  
 Db 331 TTGCAGATACAGGCAAGCTGATGACACCATGAAAGCGCATGAGCGCGCCGATGT 390  
 QY 92 GlyValThrAspThrAsnSerTyrAlaIleTrpIleGluValGlieSerAspLeuPheAla 111  
 Db 391 GGTGTTCCAGACAG-----TTTGGG 411  
 QY 112 ArgHisArgThrLysMetCysArgLysLysArgPheAlaLysGlnGlyAsnLysTrpTyr 131  
 Db 412 GCTGAGATCAAGGCCCATGTTTGAAGAAAGCGCTACGCCATCCAGGTTCTCAATGGCAA 471  
 QY 132 LysGlnHisLeuSerTyrArgLeuValAsnTrp-ProGlnHisLeuArgSerArgGlnPhe 151  
 Db 472 CATATGAATTAATTAATTTCTGTCATCAGAAATTACACCCCAAGTGGCGAGTATGCCACA 531  
 QY 151 eGlyAlaProCysAlaProProSerSerCysGlyAlaThrSerGlnAArgTrpSerSerGly 171  
 Db 532 TACGAGGCGCATTCGCAAGCGCTTCGCGGTGGAGAGTCCACACCATCGCTTCGCC 591  
 QY 171 YArg---ProGlnProGlnAlaProLeuThrSerGlySerPro-----SerSe 186  
 Db 592 GAGGTGCTCATGCTATGCTATCCGTAGAGGCCATGAGAGCGCCACATCATGATCTTC 651  
 QY 186 rLysGlyThrThrMetGlyTrpAlaMet-ProLeuMetAlaGlnGlyAlaProTrpArg 206  
 Db 652 TTTCGCCAGGGCTTCATGGCGACAGACCCCTTCGATGTCGAGGGCGGCTTCGCGCC 711  
 QY 206 gThrProPheLeuProArg-----ArgGlyGlnAlaHisPheAspGlnAspGlyUar 223  
 Db 712 CATGCCAC-TTCCAGGGCCCAACATGGAGGAGACACCACTTGACTTCGCCAGGCC 770  
 QY 223 gTrpSerLeuSerArgArg-----ArgGlyArgAsnLeuPheValValLeuAlaHisGly 241  
 Db 771 TTGGACTGTCAGGAATGAGATGTGAAATGAAATGACATCTTCGTGGCTGTGCACGA 830  
 QY 241 uIleGlyHisThrLeuGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaProTyr 261  
 Db 831 GCTGGGCGCATGCTCGGGCTCGAGCATTCAGTGACCCCTCGGCATCATGCACTT 890  
 QY 261 rTyrLysArgLeuGly---ArgAspAlaLeuLeuSerTrpAspAspValLeuAlaValGly 280  
 Db 891 TTACCACTGATGACACGAGAAATTTGTGCTTCGATGATGACCCGCGGGCATCCA 950  
 QY 280 nSerLeuTyrGlyLysProLeuGlyGlySerValAlaValGlnLeuProGlyLysLeuPhe 300  
 Db 951 GCACCTTATGGGGTGAAGTCAGG-----TTCCCCACCAAGATGCC 992  
 QY 300 ethrAspPheGlnuThrTrpAspSerTyrSer---ProGlnGlyArgArgProGlnuThrGly 319  
 Db 993 CCTCAACCCAGACTACCTCCCGGCTTCGTGTCCTGATTAACCAAAAACCCCACTTA 1052  
 QY 319 nGlyProLysTyrCysHisSerSerPheAspAlaIleThrValAspArgGlnGlnLeu 339  
 Db 1053 TGGGCGCAACATCTGTGACGGGAATTTGACACCGTCGACCTGCCA---GGGAGAT 1109  
 QY 339 uTyrIlePheLysGlySerHisPheTrpGluVal-----AlaAlaAspGlyAs 355  
 Db 1110 GTTGTCTTCAAGAGCGCTGCTGCGGGTGAGGAATACCAAGTATGATGGA-- 1167  
 QY 355 nValSerGluProArgProLeuGlnGlyUarGTrpValGlyLeuProProAsnIleGlyAla 375  
 Db 1168 -----TACCAATGCCCATTTGGCCAGTTCTGGCGGGCGCTGCTGCCATCAAC-- 1218  
 QY 375 aAlaAlaValSerLeuAsnAspGlyAspPheTyrPhePheLysGlyGlyArgCysTrpArg 395  
 Db 1219 -ACTGCTACAGAGAGAAAGATGGCAAAATTCGTCTTCTCAAGAGACACAGCATTTGGGT 1277

QY 395 gPheArgGlyProLysProValTrpGlyLeuProGlnLeuCysArg-----AlaGlyGly 413  
 Db 1278 GTTGTATGAGGGGTCCCTGGAAACCTGCTACCCCAAGCAATTAAAGAGCTGGGCGGAG 1337  
 QY 413 yLeuProArgHisPro---AspAlaAlaLeuPhePheProProLeuArgArgLeuIleLe 432  
 Db 1338 GCTGCTTCCGACAAAGATTTGATGCTGCTCTCTTCATGATGCCAATGGAAGACCTACTT 1397  
 QY 432 uPheLysGlyAlaArgTyrTyrValLeuAlaArgGlyGlyLeuGlnValGluProTyrTyr 452  
 Db 1398 CTTCCGTGGAACAAGTACTTACCTGTTTCAACGAAGAGCTCAGGCGACGTATGCGAGTA 1457  
 QY 452 rProArgSerLeuGlnAspTrpGlyGlyIleProGlnGlnValSerGlyAlaLeuProArg 472  
 Db 1458 CCCCAAGAACATCAAAAGTCTGGAAGGGATCCAGATCTCCAGAGGGTCAATTCATGGG 1517  
 QY 472 gProAspGlySerIle---IlePhePheArgAspAspArgTyrTrpArgLeuAspGlnAl 491  
 Db 1518 CAGCATGAAGTCTTCACTTACTTCTTCAAGGGGACAAATTAATTCGAAATTCACAAACA 1577  
 QY 491 aLysLeuGlnAlaThrThrSerGlyArgTrpAlaThrGluLeuProTrpMetGlyCys 510  
 Db 1578 GAAGCTGAAGGTAGAACCGGGCTACCCCAAGTACGCCCTGAGGAACTGGATGGGCTGC 1635

Search completed: June 14, 2005, 11:31:25  
 Job time : 303 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 14, 2005, 05:24:43 ; Search time 6714 Seconds

(without alignments)  
3752.862 Million cell updates/sec

Title: US-10-791-980-6  
Perfect score: 2834  
Sequence: 1 MVARGLLRALQLLMGL.....WATELPMWCHWANSGLAF 520

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-MODEL=frame+2n.model -DEV=xlh  
-O=/sgn2.1/USPTO/seqpool/US10791980/runat\_13062005\_104718\_19915/app\_query.fasta\_1.711  
-DB=genemb1 -QFMT=faasap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10791980 @CCN 1.1 4306 @runat\_13062005\_104718\_19915 -NCPU=6 -ICPU=3  
-NO\_MMAP -LRGEBUFFER -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Genemb1: \*  
1: gb\_ba: \*  
2: gb\_hg: \*  
3: gb\_in: \*  
4: gb\_cm: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_srs: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 2834  | 100.0       | 1597   | 6     | AR534532 Sequence |
| 2          | 2834  | 100.0       | 1597   | 6     | AX358477 Sequence |
| 3          | 2792  | 98.5        | 1560   | 6     | AX398367 Sequence |
| 4          | 2792  | 98.5        | 1563   | 6     | AX206683 Sequence |

|    |        |      |        |    |                     |
|----|--------|------|--------|----|---------------------|
| 5  | 2792   | 98.5 | 1817   | 9  | AF219624 Homo sapi  |
| 6  | 2792   | 98.5 | 1985   | 6  | AR528524 Sequence   |
| 7  | 2792   | 98.5 | 1985   | 6  | AX464010 Sequence   |
| 8  | 2792   | 98.5 | 1985   | 6  | AY358987 Homo sapi  |
| 9  | 2792   | 98.5 | 2241   | 6  | AX083147 Sequence   |
| 10 | 2792   | 98.5 | 2484   | 9  | AF330002 Homo sapi  |
| 11 | 2792   | 98.5 | 2527   | 6  | AX398365 Sequence   |
| 12 | 2784   | 98.2 | 1753   | 9  | AF315683 Homo sapi  |
| 13 | 2771.5 | 97.8 | 2091   | 6  | BC002631 Homo sapi  |
| 14 | 2764   | 97.5 | 2275   | 6  | AR211784 Sequence   |
| 15 | 2721   | 96.0 | 2336   | 6  | CO842403 Sequence   |
| 16 | 2721   | 96.0 | 2336   | 6  | AK122604 Homo sapi  |
| 17 | 2324.5 | 82.0 | 1335   | 6  | AX398377 Sequence   |
| 18 | 2324.5 | 82.0 | 2310   | 6  | AX398375 Sequence   |
| 19 | 2261   | 79.8 | 2340   | 10 | AY065653 Mus muscu  |
| 20 | 2213   | 78.1 | 2310   | 10 | AY071828 Mus muscu  |
| 21 | 2142   | 75.6 | 2268   | 10 | AY071829 Mus muscu  |
| 22 | 1217.5 | 43.0 | 156223 | 10 | AC119615 Homo sapi  |
| 23 | 1204   | 42.5 | 1916   | 5  | BC061659 Homo sapi  |
| 24 | 1198.5 | 42.3 | 86442  | 10 | AL663096 Mouse DNA  |
| 25 | 1198.5 | 42.3 | 205701 | 10 | CNS08CA6 Homo sapi  |
| 26 | 1198.5 | 42.3 | 228907 | 10 | AC124036 Homo sapi  |
| 27 | 1187   | 41.9 | 82166  | 9  | AC006237 Homo sapi  |
| 28 | 1187   | 41.9 | 179966 | 2  | AC023247 Homo sapi  |
| 29 | 1187   | 41.9 | 205040 | 9  | AC015849 Homo sapi  |
| 30 | 917    | 28.8 | 471    | 6  | CQ725232 Sequence   |
| 31 | 654    | 23.1 | 953    | 3  | BC011774 Homo sapi  |
| 32 | 617    | 21.8 | 368    | 6  | AX262107 Sequence   |
| 33 | 610.5  | 21.5 | 2315   | 3  | AF271666 Drosophila |
| 34 | 610.5  | 21.5 | 3116   | 3  | AY118405 Homo sapi  |
| 35 | 605    | 21.3 | 2323   | 5  | AB047651 Oryzias    |
| 36 | 602    | 21.2 | 1707   | 6  | CO581991 Sequence   |
| 37 | 602    | 21.2 | 2010   | 6  | AX481490 Sequence   |
| 38 | 602    | 21.2 | 3513   | 6  | CO715551 Sequence   |
| 39 | 602    | 21.2 | 3530   | 6  | A46722 Sequence     |
| 40 | 602    | 21.2 | 3530   | 6  | AR110168 Sequence   |
| 41 | 602    | 21.2 | 3530   | 6  | AR211699 Sequence   |
| 42 | 602    | 21.2 | 3530   | 9  | Z48482 H. sapiens m |
| 43 | 602    | 21.2 | 4263   | 9  | BC055428 Homo sapi  |
| 44 | 601.5  | 21.2 | 3517   | 5  | AY324196 Danto rer  |
| 45 | 599    | 21.1 | 3458   | 9  | BC006412 Homo sapi  |

## ALIGNMENTS

RESULT 1  
AR534532  
LOCUS AR534532 1597 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 3 from patent US 6734005.  
ACCESSION AR534532  
VERSION AR534532.1 GI:53924838  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1597)  
AUTHORS Wood T., Ekblom J., Holmgren E. and Kihlen M.  
TITLE Matrix metalloproteinases  
JOURNAL Patent: US 6734005-A 3 11-MAY-2004;  
FEATURES  
source Location/Qualifiers  
1..1597  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,936-113  
Score: 2834.00 Length: 1597  
Percent Similarity: 100.00% Matches: 520  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-791-980-6 (1-520) x AR534532 (1-1597)

```

QY      1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHisLeu 20
Db      35 ATGGTCGGCGCGCTGCGCTCTGCTCGCGCCCTCGAGCTGCTACTGTGGGCCCACTTG 94
QY      21 AspAlaGlnProAlaGlnArgGlyGlyGlnGlnLeuArgGlySerGlyLysAlaGlnAlaPheLeu 40
Db      95 GACCCCAAGCCCGCGAGCGCGGAGCCAGAGACTCGCGCAAGAGAGCGAGGACCTTCTCA 154
QY      41 GlnLysTrpGlyTrpLeuAangGlnGlnValProLysAlaProThSerThrArgPheSer 60
Db      155 GAGAACTACGGAATACCTCAATGAACAGCTCCCAAGCTCCCACTCCACTCGATTGACG 214
QY      61 AspAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuAspArg 80
Db      215 GATGCCATTCAGAGCGTTTCACTGGGTGCTCCAGCTACCTGTGAGGGGCTTTGACCGC 274
QY      81 AlAThLeuArgGlnMetThrArgProArgGlyValThrAspThrAsnSerTyrAla 100
Db      275 GCCACCTCGCGCGAGATGACTGCTCCCGCTGCGGGGGTTACAGATACCAAGATTATGCG 334
QY      101 AlaTrpAlaGlnArgIleSerAspLeuPheAlaArgHisArgThrLysMetArgArgLys 120
Db      335 GCTGGGCTGAGAGAGATGATGACTTGTGTCTAGACACCGGACCAAAATGAGGCGTAAAG 394
QY      121 LysArgPheAlaLysGlnGlyAsnLysTrpLysGlnHisLeuSerTyrArgLeuVal 140
Db      395 AAACCTTTGGCAAGCAAGCAAGTAACTGATGACAGACCACTCTCTCAACCGCTGCTG 454
QY      141 AsnTrpProGlnHisLeuArgSerArgGlnPheGlyAlaProCysAlaProPheSerSer 160
Db      455 AACTGGCTTGAGCAATCTCCGAGCGCGAGCTGCGGCGCGCGTCCGCGCCCTTCCAGT 514
QY      161 CysGlyAlaThrSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuThr 180
Db      515 TGTGAGCAACAGTCTCAGCGCTGAGATTCTGAGAGGCCCAAGCAAGCGCCCGCTGACA 574
QY      181 SerGlySerProSerSerLysGlyThrThrThrMetGlyTrpAlaMetProLeuMetAla 200
Db      575 TCCGGCTCACCCTTCTTCCAAAGGAGCAACAGATGGCTGGCAATCCTTTGATGGCC 634
QY      201 GlnGlyAlaLeuProTrpArgTrpProPheLeuProArgArgGlyGlnAlaHisPheAspGln 220
Db      635 CAGGGGGCGCGCTGGCGCAAGCTTCTTCCGCCCGCGCGGCAAGCGCACTTGCAACAA 694
QY      221 AspGlnArgTrpSerLeuSerArgArgArgGlyArgLeuLeuPheValIleuAlaHis 240
Db      695 GATGAGCGCTGGTCCCTGAGCGCGCGCGGCAACCTGTTCTGTGTGGTGGCGAC 754
QY      241 GlnIleGlyHisThrLeuGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaPro 260
Db      755 GAGATCGGTCAACAGCTTGGCTCAACCACTGCGCGCGCGCGCGCTCATGGCGCC 814
QY      261 TyrTrpLysArgLeuGlyArgAspAlaLeuLeuSerTrpAspArgValIleuAlaValGln 280
Db      815 TACTACAAAGAGCTGGGCGCGGAGCGCGCTGCTGAGGAGCAACCTGCTGGCGCGG 874
QY      281 SerLeuTrpGlyLysProLeuGlyGlySerValAlaValGlnLeuProGlyLysLeuPhe 300
Db      875 AGCTGTATGGAGAGCCCTTAGGGGCTCAGTGGCGCTCAGAGTCCAGAGAAAGCTGTTT 934
QY      301 ThrAspPheGlnTrpAspSerTyrSerProGlnGlyArgArgProGlnLysThrGlnLys 320
Db      935 ACTGACTTTAGACCTGGGACTCTCAAGCCCAAGAGAGCGCGCTGAAACGAGAGGC 994
QY      321 ProLysTrpCysHisSerSerPheAspAlaIleThrValAspArgGlnGlnLeuLys 340
Db      995 CCTAAATATCGCACCTTCTTCAATGCAATCACTGATGACAGGCAACACACTGTAC 1054
QY      341 IlePheLysGlySerHisPheTrpGlnValAlaIleAspGlyAsnValSerGlnProArg 360
Db      1055 ATTTTAAAGGAGCAATTTCTGGAGGTGGAGCTGATGCAACGTTCTCAGAGCCCGT 1114

```

```

QY      361 ProLeuGlnGlnArgTrpValGlyLeuProProPheAlaIleGlnAlaIleValSerLeu 380
Db      1115 CCAGTCCAGAGAAAGATGGGTGGGCTGCCCGCCCAACATTGAGGCTGCGGCAGTGTCAATTG 1174
QY      381 AsnAspGlyAspPheTrpPheLysGlyGlyArgCysTrpArgPheArgGlyProLys 400
Db      1175 AATGATGAGATTTTCTACTTCTTCAAGGGGGTGAATGCTGAAGTTCCGGGGCCCCAAG 1234
QY      401 ProValTrpGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHisProAspAla 420
Db      1235 CCAGTGGGGGTCTCCCAAGACTGTCCGCGGAGAGGGGCGCTGCGCCCGCTCGAGCGC 1294
QY      421 AlaLeuPhePheProProLeuArgArgLeuIleLeuPheLysGlyAlaArgTyrTrpVal 440
Db      1295 GCCCTTCTTCTCTCTCTGCGCGCGCTCATCTCTTCAAGGGGTCCCGCTCACTAGTG 1354
QY      441 LeuAlaArgGlyGlyLeuGlnValGluProTyrTrpProArgSerLeuGlnAspTrpGly 460
Db      1355 CTGGCCCGAGGGGAGCTGCAAGTGAAGCCCTTACTACCCCGCAAGTCTGCAGAGCTGGGA 1414
QY      461 GlyIleProGlnGlnValSerGlyAlaLeuProArgProAspGlySerIleIlePhePhe 480
Db      1415 GGCATCCTTGAAGAGGTCAAGCGCGCGCTGCGAGGCGGAGTCTCATCATCTTCTTC 1474
QY      481 ArgAspAspArgTrpTrpArgLeuAspGlnAlaLysLeuGlnAlaThrThrSerGlyArg 500
Db      1475 CGAGATACCGCTACTGCGGCTTCAACAGGCCCAACTGCAAGGCAACCACTCGGGCGCG 1534
QY      501 TrpAlaThrGlnLeuProTrpMetGlyCysTrpHisAlaAsnSerGlySerAlaLeuPhe 520
Db      1535 TGGGCCCAACGAGCTCCCTGATGAGGTGTGCTGATGCACTGGGGAGCGCCCTGTTC 1594

```

# RESULT 2

```

AX358477 1597 bp DNA linear PAT 13-FEB-2002
LOCUS AX358477
DEFINITION Sequence 3 from Patent WO0190326.
ACCESSION AX358477
VERSION AX358477.1 GI:18675088
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Holmgren, E., Kihlen, M., Wood, T. and Ekblom, J.
TITLE Novel matrix metalloproteinases
JOURNAL Patent: WO 0190326-A 3 29-NOV-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES
Source location/Qualifiers
1..1597
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1..936-113
Score: 2834.00 Length: 1597
Percent Similarity: 100.00% Matches: 520
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
DB: Indels: 0
Gaps: 0

```

```

US-10-791-980-6 (1-520) x AX358477 (1-1597)
QY      1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHisLeu 20
Db      35 ATGGTCGGCGCGCTGCGCTCTGCTCGCGCCCTCGAGCTGCTACTGTGGGCCCACTTG 94
QY      21 AspAlaGlnProAlaGlnArgGlyGlyGlnGlnLeuArgGlySerGlyLysAlaGlnAlaPheLeu 40
Db      95 GACCCCAAGCCCGCGAGCGCGGAGCCAGAGCTCGCGCAAGAGCGGAGGCAATTCCTTA 154

```

Oy 41 GluYbTYRGIYTYrLeuAaNGluGlnValProLYaAlaProThrSerThzArgPheSer 60  
 Db 155 GAGAAATGACGATACCTCAATGAACAAGGATCCCAAGCTCCACTGATTCAGC 214  
 Oy 61 AspaAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuAaArg 80  
 Db 215 GATGCATACAGAGCTTTCAGTGGGTGTCCAGCTACCTGACGGGGGTGTGAGCCG 274  
 Oy 81 AlaThrLeuArgGlnMetThzArgProArgCysGlyValTrpAspThzAsnSerTrpAla 100  
 Db 275 GCCACCTGGCCCAATGATCTGCTCCGCTGCGGGATTACAGATACCAAGTTATGCG 334  
 Oy 101 AlaTrpAlaGluArgIleSerAspLeuPheAlaArgHISArgThzLysMetArgArgLys 120  
 Db 335 GCTCGGGCTGAGAGATCAGTGACTTTGTTGTACACCCGGAACAAATAGAGCGGTAG 394  
 Oy 121 LysArgPheAlaLysGlnGlyAaLysLysTrpTyrLysGlnHISLeuSerTrpArgLeuVal 140  
 Db 395 AAACCTTTGGCAACCAAGGTAAATGTACAAAGCAGCACTCTCTACCGCTGGTG 454  
 Oy 141 AsnTrpProGlnHISLeuArgSerArgGlnPheGlyAlaProCysAlaProProSerSer 160  
 Db 455 AACTGGCTGAGCATCTCCGAGCCGAGTCGGGGCGCGCTGGCGCGCTTCCAGT 514  
 Oy 161 CysGlyAlaTrpSerSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuThr 180  
 Db 515 TGTGAGAGCAGCTCCAGCGCTGAGATTCTGGAGAGCCCAAGCCCAAGCCCGCTGACA 574  
 Oy 181 SerGlySerProSerSerLysGlyThrThzMetGlyTrpAlaMetProLeuMetAla 200  
 Db 575 TCCCGCTCAGCTTCTTCAAGGGGAGCAACAAGATGGGCTGGGCAATCCTTGTATGCGC 634  
 Oy 201 GlnGlyAlaProTrpArgTrpPheLeuProArgArgGlyGlnAlaHISpHeAspGln 220  
 Db 635 CAGGGGGGCGCTGGCGGACCGCTTCTGCCCCCGGGGAGAGCGACTTCACACAA 694  
 Oy 221 AspGluArgTrpSerLeuSerArgArgArgGlyArgAaLeuPheValLeuAlaHIS 240  
 Db 695 GATGAGCGCTGGTCTCTGAGCGCGCGCGGCGGCAACTGTTGTGTGTGGCGGAC 754  
 Oy 241 GluIleGlyHISThrLeuGlyLeuThrHISerProAlaProArgAlaLeuMetAlaPro 260  
 Db 755 GAGATCGGTCACACCTTGCGCTCACCACTGCGCGCGCGCGCTCATGCGCGCC 814  
 Oy 261 TyrTyrLysArgLeuGlyArgAspAlaLeuLeuSerTrpAspArgValLeuAlaGln 280  
 Db 815 TACTACAAAGAGGCTGGCGCGGACGCTGCTCAGCTGGGACGAGCTGCGCTGCGCAG 874  
 Oy 281 SerLeuTyrGlyLysProLeuGlyGlySerValAlaValGlnLeuProGlyLysLeuPhe 300  
 Db 875 AGCGCTGATGGAGAACCCCTAGGGGGCTCAGTGGCGCTCAGCTCCAGGAAGCTGTTC 934  
 Oy 301 ThrAspPheGluThrTrpAspSerTyrSerProGlnGlyArgArgProGlnThrGlnGly 320  
 Db 935 ACTGACTTTGAGACTCGGGACTCTACAGCCCAAGGAGGCGCGCTGAAGACGAGGGC 994  
 Oy 321 ProLYbTYRGIYbHISerSerPheAspAlaIleThrValAspArgGlnGlnGlnLeuTyr 340  
 Db 995 CTTAAATCTGCGACTTCTCTTCATGCGACATGTAAGACAGGCAACAGCACTGTAC 1054  
 Oy 341 IlePheLysGlySerHISpHeTrpGlyValAlaAlaAspGlyAspValSerGlnProArg 360  
 Db 1055 ATTTTAAAGGAGCATTCTGGAGGCTGGAGCTGAGTGGCAAGCTCTCAGAGCCCT 1114  
 Oy 361 ProLeuGlnGluArgTrpValGlyLeuProProAsnIleGluAlaAlaValSerLeu 380  
 Db 1115 CCACTGAGCAAAAGATGGGTGGGCTGCCCAACATTGAGGCTGGCGGAGTGCATTG 1174  
 Oy 381 AsnAspGlyAspPheTrpPhePheLysGlyGlyArgCysTrpArgPheArgGlyProLys 400  
 Db 1175 AATGATGAGATTTACTTCTTCAAAAGGGGTGATGCTGAGGGTCCGGGGCCCAAG 1234

Oy 401 ProValTrpGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHISProAspAla 420  
 Db 1235 CCAATGTGGGGTCTCCCAAGACTGTGCGGGCAGGGGGCTGCCCGCATCTGACGCC 1294  
 Oy 421 AlaLeuPhePheProProLeuArgArgLeuIleLeuPheLysGlyAlaArgTyrTrpAla 440  
 Db 1295 GCCCTCTTCTCCCTCTGCGCGCTCATCTCTCAAGGGGTGCCGCTACTACAGCTG 1354  
 Oy 441 LeuAlaArgGlyGlyLeuGlnValGluProTyrTrpProArgSerLeuGlnAspTrpGly 460  
 Db 1355 CTGGCCCGAGGGGAGACTCCAAAGTGAAGCCCTACACCCCAATGTGACAGACTGGGA 1414  
 Oy 461 GlyIleProGlnGluValSerGlyAlaLeuProArgProAspGlySerIleIlePhePhe 480  
 Db 1415 GGCAATCCCTGAGAGGTCAAGCGCGCTGCCGAGAGCCGATGCTCATCATCTTCTTC 1474  
 Oy 481 ArgPheAspArgTyrTrpArgLeuAspGlnAlaLysLeuGlnAlaTrpThzSerGlyArg 500  
 Db 1475 CGAATGACCGCTACTGCGCTGACCAAGCCAAACTGCAAGGCAACCACTCGGGCGC 1534  
 Oy 501 TrpAlaThrGluLeuProTrpMetGlyCysTrpHISAlaAsnSerGlySerAlaLeuPhe 520  
 Db 1535 TGGCCACCGAGCTGCCCTGATGGGCTGTGCGATGCACTCGGGAGAGCGCTTTC 1594

**RESULT 3**  
 AX398367 1560 bp DNA linear PAT 27-MAY-2002  
 LOCUS AX398367  
 DEFINITION Sequence 3 from Patent WO0220739.  
 ACCESSION AX398367  
 VERSION AX398367.1 GI:21261134  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**REFERENCE**  
 1  
 CURTIS, R.A.  
 TITLE Human matrix metalloproteinase  
 JOURNAL Patent: WO 0220739-A 3 14-MAR-2002;  
 Millennium Pharmaceuticals, Inc. (US)

**FEATURES**  
 SOURCE  
 1..1560  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 1,2e-111 Length: 1560  
 Score: 2792.00 Matches: 519  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 98.52% Indels: 2  
 DB: Gaps: 0

US-10-791-980-6 (1-520) x AX398367 (1-1560)

Oy 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHISLeu 20  
 Db 1 ATGGTGGGCGCGTCTCTCTGCGCGCTTGGACGCTGACGTGGGGCACCTG 60  
 Oy 21 AspAlaGlnProAlaGluArgGlyGlyGlnIleuAspArgLysAlaGluAlaPheLeu 40  
 Db 61 GACCCCAAGCCCGGAGCGCGGAGCCAGAGACTGGCAAGAGCGGAGGCACTTCTTA 120  
 Oy 41 GluYbTYRGIYTYrLeuAaNGluGlnValProLYaAlaProThrSerThzArgPheSer 60  
 Db 121 GAGAAATGACGATACCTCAATGAACAAGGATCCCAAGCTCCACTGATTCAGC 180  
 Oy 61 AspaAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuAaArg 80  
 Db 181 GATGCATACAGAGCTTTCAGTGGGTGTCCAGCTACCTGACGGGGGTGTGAGCCG 240

QY 81 AlAThPLeuAArgInWetThrAProArgSGIValThrAspThrAsnSerTYrAla 100  
 DB 241 GCCACCTCTGGCCAGATGACTCTCCCGCTCGGGGTTACAGATACCAACAGTTATGCG 300  
 QY 101 AlATrPAAGluArgIleSerAspLeuPheAlaArgHISArgThyIysMetArgArgIys 120  
 DB 301 GCGTGGGCTGAGAGATCTAGATCTGTGTCTGACACCCGACCAAAATGAGCGTTAG 360  
 QY 121 LysArgPheAlaLysGlnGlyValAsnLysTrpTyrLysGlnHISLeuSerTyrAlaGleVal 140  
 DB 361 AAACGCTTTGGCAAAACAGGTAAACAAATGTACAAAGACACCTCTCTTACCGCTGTGT 420  
 QY 141 AsnTrpProGluHISLeu-ArgSerArgGlnPheGlyAlaProCysAlaProProSerSe 160  
 DB 421 AACTGCGCTGAGCATCTGCGGACCCGACAGTTGCGGGCGCCCTGTGCGCGCTTCCAG 480  
 QY 160 TCGAGIValThrSerGlnHISArgTrpSerSerGlyArgProGlnProGlnAlaProLeuTh 180  
 DB 481 TTGTGAGACAAAGCTTCTAGCGCTGAGATTCTGGAGGCCCCAGCCACAGGCCCGCTGAC 540  
 QY 180 rSerGlySerProSerSerIysGlyThrThrMetGlyTTPAlaMetProLeuMetAl 200  
 DB 541 ATCCGGCTCACCTTCTTCCAAAGGAGCAACAAAGATGGCTGGCAATGCTTTGATGCG 600  
 QY 200 AGInGlyAlaProTrpArgThrProPheLeuProArgArgIysGlnAlaHISpHeArgI 220  
 DB 601 CCAAGGGGCGCCCTGGCGACCCCTTC-CTGCGCGCGCGGCGCAAGCGCACTTCCACA 659  
 QY 220 nAspGluArgTrpSerLeuSerArgArgArgIysArgAsnLeuPheValIleuAlaHIS 240  
 DB 660 AGATGAGCCCTGCTGCTGAGCCCGCGCGGCGCAACCTGTCTGTGTCTGCGCA 719  
 QY 240 AGIuIleGlyHISThrLeuGlyLeuThrHISerProAlaProArgAlaLeuMetAlaPr 260  
 DB 720 CCAAGATCGGTACACGCTTGGCTTCCACCCTCGCCCGCGCGCGCTCATGCGCGC 779  
 QY 260 CTyTrpTyrArgArgLeuGlyArgAspAlaLeuLeuSerTrpAspAspValIleuAlaVal 280  
 DB 780 CTACTACAAAGAGCTGGGCGCGCAACGCGCTCTCACTGAGACGAGCTGCGCGCTGCA 839  
 QY 280 nSerLeuTyrGlyLysArgProLeuGlyLysSerValAlaValGlnLeuProGlyLysLeuPh 300  
 DB 840 GAGCTGTATGGAGAGCCCTTAAAGGAGCTCACTGCGCGCTCCAGCTCCAGGAAAGCTGT 899  
 QY 300 eThrAspPheGlnThrTrpAspSerTyrSerProGlnGlyArgArgProGluThrGlnG 320  
 DB 900 CACTGACTTTGAGACCTGGGACTCTTACAGCCCCCAAGAGGCGCTTAAACGACAGG 959  
 QY 320 YProLysTyrCysHISerSerPheAspAlaIleThrValAspArgGlnGlnLeuTyr 340  
 DB 960 CCTTAATAGTCACTCTTCTGATGCACTAGCTGTAGACAGGCAACAGCACTGTA 1019  
 QY 340 rIlePheLysGlySerHISpHeTrpGluValAlaAlaAspGlyAsnValSerGluProAr 360  
 DB 1020 CATTTTAAAGGAGGACATTTCTGGAGGTGGACGCTGAGGCAAGCTCTCAGAACCCCG 1079  
 QY 360 gProLeuGlnIuArgTrpValGlyLeuProProAsnIleGluAlaAlaValSerLe 380  
 DB 1080 TCCACTGCAAGAAAGATGGGTGGGCTGCCCCCAACATTGAAGGTGGCGGAGTCTATT 1139  
 QY 380 uAsnAspGlyAspPheTyrPhePheLysGlyArgCysTrpAspPheArgIysProLy 400  
 DB 1140 GAATGATGAGATTTCTACTTCTTCAAAAGGGGTGAGTCTGAGAGTTTCCGGGCCCCAA 1199  
 QY 400 gProValTrpGlyLeuProGlnLeuCysArgAlaGlyLysLeuProArgHISProAspAl 420  
 DB 1200 GCCAGTGTGGGTCTCCACAGCTGTGCGGCAAGGGGCGCTGCGCGCATCTGACGCG 1259  
 QY 420 aAlaLeuPhePheProProLeuHISArgLeuIleuPheLysGlyAlaArgTyrTyrVal 440  
 DB 1260 CGCCCTCTTCTCCCTCTGCGCGCGCTCATCTTCAAGGTTCCCGCTACTACAT 1319  
 QY 440 IleuAlaArgGlyGlyLeuGlnValGluProTyrTyrProArgSerLeuGlnAspTrpG 460

DB 1320 GCTGCGCCGAGGGGAGCTGAGATGAGGCCCTTACTACCCCGAAGTCTGCAGAGCTGGG 1379  
 QY 460 YGlyIleProGlnGlnValSerGlyAlaLeuProArgProArgIysSerIleIlePhePh 480  
 DB 1380 AGGATCTCCCTGAGAGAGTCTAGCGGCGCCCTGCGAGGCCCGCATGCTCATCTTCTT 1439  
 QY 480 eArgAspAspArgTyrTrpArgLeuAspGlnAlaLysLeuGlnAlaThrThrSerGlyAr 500  
 DB 1440 CCGAGATGACCGCTACTCTGCGCTTGCACAGGCCCAACTCAGGCAACCACTCGGGCG 1499  
 QY 500 gTrpAlaThrGlyLeuLeuProTrpMetGlyCysTrpHISAlaAsnSerGlySerAlaLeuPh 520  
 DB 1500 CTGGGCGACCGAGCTGCGCTGATGAGGCTGCTGCGCATGCCAATCTCGGAGAGCGCTGTT 1559  
 QY 520 e 520  
 DB 1560 C 1560

RESULT 4  
 AX206683 1563 bp DNA linear PAT 30-AUG-2001  
 LOCUS AX206683  
 DEFINITION Sequence 1 from Patent WO0155428.  
 ACCESSION AX206683  
 VERSION AX206683.1 GI:15394591  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Souhan, C.D. and Hughes, A.S.  
 TITLE Mproct45 metalloprotease  
 JOURNAL Patent: WO 0155428-A 1 02-AUG-2001;  
 SmithKline Beecham P.L.C. (GB)  
 FEATURES  
 source  
 Location/Qualifiers  
 1..1563  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

Alignment Scores:  
 Pred. No.: 1,2e-111 Length: 1563  
 Score: 2792.00 Matches: 519  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatch: 1  
 Query Match: 98.52% Indels: 2  
 DB: 6 Gaps: 0

US-10-791-980-6 (1-520) x AX206683 (1-1563)  
 QY 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHISLeu 20  
 DB 1 ARGTRCGCGCGGTGGCTCTCTCTGCGCGCTGACAGTGTCTACTGTGGGCCCACTTG 60  
 QY 21 AspAlaGlnProAlaGluArgGlyGlyGlnGluLeuArgLysGlnAlaLysLeu 40  
 DB 61 GAGGCCAGCCCGGAGCGCGGAGGCGCAGAGCTCGCAAGAGCGGAGGATTTCTTA 120  
 QY 41 GluLysTyrGlyTyrLeuAsnGlnGlnValProLysAlaProThrSerThrArgPheSer 60  
 DB 121 GAGAAATACAGATTCATGACATGAAAGGTCCTCCAAAGCTCCACCTCCACTCGATTACG 180  
 QY 61 AspAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValIleuAspArg 80  
 DB 181 GATGCCATACAGAGCGTTTCTAGTGGGTGCCAGCTACTGTCAAGCGGTGTGACCGC 240  
 QY 81 AlAThLeuArgGlnMetThrArgProArgCysGlyValThrAspThrAsnSerTYrAla 100  
 DB 241 GCCACCTCTGGCCAGATGACTCTCCGCTGCGGGGTTCAGATACCAACAGTTATGCG 300  
 QY 101 AlATrPAAGluArgIleSerAspLeuPheAlaArgHISArgThrIysMetArgArgIys 120

```
Db 301 GCCTGGGCTGAGAGATCATGACTGTTTGTCTACACACCGGACCAAAATGAGGGGTAG 360
Qy 121 LysArgPheAlaLysGlnGlyValAsnLysTrpTyrLysGlnNH1sLeuSerTyrArgLeuVal 140
Db 361 AAAGCCTTTCAGAACAGAGTAAACAATGTATCAAGACAGACTCTCTCCATCGGCTGGG 420
Qy 141 AsnTrpProGlnNH1sLeu-ArgSerArgGlnPheGlyValProCysValProProSerSe 160
Db 421 AACTGGCCTGAGCATCTGCCGAGCGCGGAGTTCGGGGCGCGTCCGGCGCGCTTCAG 480
Qy 160 rCysGlyValArhSerGlnA-gTrpSerSerGlyValArgProGlnProGlnAlaProLeuTh 180
Db 481 TTGTGGAGCAACGTCTCAGCGCTGAGGTCTGGGAGCGCCCAAGCCAGCGCCCTGAC 540
Qy 180 rSerGlySerProSerSerLysGlyTThrThrMetGlyTThrAlaMetProLeuMetAla 200
Db 541 ATCCGGCTACCTTCTTCAGAGGGGACCAACAGATGGGCTGGGGCAATGCTTGTATGC 600
Qy 200 aGlnGlyValArProTrpArgTrhProPheLeuProArgArgGlyGlnAlaNH1sPheAspG 220
Db 601 CCAAGGGGCGCGCTGGCGGACGCTTC-CTGGCGCGCGCGCGGAGGCACTTCGACCA 659
Qy 220 nAspGluArgTrpSerLysSerArgArgArgGlyValArgAsnLeuPheValValLeuAlaH 240
Db 660 AGATGAGCGCTGTGCTCCAGAGCGCGCGCGCGCGCAACCTGTTCTGTGCTGGCGCA 719
Qy 240 aGlnNH1sLeuArhTrhLeuGlyLeuThHisSerProAlaProArgAlaLeuMetAlaPr 260
Db 720 CGAGATCGGTCACAGCTTGGCTTCACCACTCGCGCGCGCGCGCGCTCATGGCC 779
Qy 260 cTyrTyrLysArgLeuGlyValArgAspAlaLeuLeuSerTrpAspAspValLeuAlaValG 280
Db 780 CTACTACAAAGAGCGCTGGCGCGGACGCTGCTGACGTCGAGCAACGTCGTGGCGCTGCA 839
Qy 280 nSerLeuTyrGlyLysProLeuGlyGlySerValAlaValGlnLeuProGlyLysLeuPh 300
Db 840 GAGCGCTGTATGGAAAGCCCTAGAGGGGCTCAGTGGCCCTCAGACTCCAGAGAAAGCTGT 899
Qy 300 eThrAspPheGlnhTrpAspSerTyrSerProGlnLysArgArgProGlnThGlnG 320
Db 900 CACTGACTTGAAGACTGGAGCTCTACAGCGCCCAAGAAAGCGCCCTGAAAGCGACGG 959
Qy 320 yProLysTyrCysNH1sSerSerPheAspAlaIleThrValAspArgGlnGlnLeuTyr 340
Db 960 CCGTAAATACAGCACTCTCTCTCGATGATCATCTTAAAGCAAGCAACGAACTGTA 1019
Qy 340 rIlePheLysGlySerHisPheTrpGlnValAlaAlaAspGlyAsnValSerGluProAr 360
Db 1020 CATTTTAAAGGAGCAATTTCTGGAGAGTGGCAGCTGATGGCAACGTCTCAGAGCCCG 1079
Qy 360 gProLeuGlnGlnLysTrpValGlyLeuProProAsnNH1sLeuAlaAlaValSerLe 380
Db 1080 TTCCTGAGAGAAAGATGGGTGGGCTGCCCCCAACATGAGGGTGGGGCATGTCTACT 1139
Qy 380 uAsnAspGlyAspPheTyrPhePheLysGlyGlyValArgCysTrpArgPheArgGlyProLy 400
Db 1140 GAATGATGAGATTTCTACTTCTTCAAAAGGGGTGAGTCTGAGAGTTCCGGGGCCCAA 1199
Qy 400 sProValTrpGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHisProAspAl 420
Db 1200 GCCAGTGGGGGTCTCCCAAGCTGTGCGGGGAGGGGGCGCTGCCCCCACTCCGAGCG 1259
Qy 420 aAlaLeuPhePheProProLeuArgArgLeuNH1sLeuPheLysGlyValArgTyrTyrVa 440
Db 1260 CGCCCTCTTCTTCTCTCTCTGCGCGCTCACTCTTCAAGGGTGGCCGCTACTACGT 1319
Qy 440 lLeuAlaArgGlyGlyLeuGlnValGluProTyrTyrProArgSerLeuGlnAspTrpG 460
Db 1320 GCTGGCCGAGGGGAGCTGCAAGTGAACCTTAACCCCGAAGTCTGCGAGACTGGGG 1379
Qy 460 yGlyTyrLeuProGluValAsnGlyValLeuProArgProAspGlySerNH1sLeuPhe 480
```

```
Db 1380 AGGCATCCCTGAGGAGGTCAAGCGCGCGCTGCCAGAGCCCGCATGAGCTTCATCTTCTT 1439
Qy 480 eArgAspAspArgGlyTyrTrpArgLeuAspGlnAlaLysLeuGlnNH1sLeuThrSerGlyAr 500
Db 1440 CCGAGATGACCGGCTTACTGCGCGCTCGACAGGCCAAACTGAGGCAACCACTCGGGCG 1499
Qy 500 gTrpAlaThrGlnLeuProTrpMetGlyCysTrpHisAlaAsnSerGlySerAlaLeuPh 520
Db 1500 CTGGGCGCACCGAGCTGCCCTGATGGGCTGTGGCATGCACTCGGGGAGCGCGCTGTT 1559
Qy 520 e 520
Db 1560 C 1560

RESULT 5
AF219624
LOCUS AF219624
DEFINITION Homo sapiens matrix metalloproteinase-28 precursor, mRNA, complete cds.
ACCESSION AF219624
VERSION AF219624.1 GI:12698337
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1817)
Loth,J., Wilson,C.L., Roby,J.D. and Parks,W.C.
Epilysin, a novel human matrix metalloproteinase (MMP-28) expressed in testis and keratinocytes and in response to injury
JOURNAL J. Biol. Chem. 276 (13), 10134-10144 (2001)
MEDLINE 21167821
PUBMED 11121398
REFERENCE 2 (bases 1 to 1817)
Loth,J.L., Wilson,C.L., Roby,J.D. and Parks,W.C.
Direct Submission
TITLE Submitted (28-DEC-1999) Pathology, Univ. Helsinki, P.O. Box 21
JOURNAL (Hartmanninkatu 3), Helsinki FIN-00014, Finland
FEATURES
source
1..1817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/cell_type="keratinocyte"
/tissue_type="testis"
1..169
170..1732
/note="neutral metalloproteinase; MMP-28; member of the neutral endopeptidase group of matrix metalloproteinases or matrixins; similar to interstitial collagenase (MMP-1)"
/codon_start=1
/product="matrix metalloproteinase-28 precursor"
/protein_id="AAK01480.1"
/db_xref="GI:12698338"
/translation="MVAVRGLILRLALQLILMGHLDAPARCGQELRKEAFLKTYG
YLNEQVKKAPTSTFSDAIRFQWVSQLPVSGLVDRATLRQMTPRCGVDTNYSAAW
AERTSDLPARRHTMRKRPAKQGNKQWYOHLSYRLVMNPEHLPEPAVGAVAPAA
LMSVSLERFWEAPATGPADIRILRTFFGDHNDIGAKFPDGGALAAAPFRRBAF
DQDERWLSRRGRNLFVLAHEITGLGTHSPAPALMAFYKRLGRDLLSWDIV
LAVSLGKPLGGGSVAQVLPKGLTDETDYSYSPQGRBETQPKYCHSFDATVD
RQOOLYIFKSHFWFVADANVSEPRPLQGRWVGLPNIILAAVSLNDGDFYFPKGR
CMRPRGPKPVWGLPOLCRAGLPHRPAALFFPLRLILFKGARYVVLARGGLQVR
YPSRLQDKGIPBEVSGALPRPDGSIIFPRDRYWRLLDQAKLQATTSGRWATBLPMK
GCHWANGSLP"
170..235
236..1729
/note="encodes matrix metalloproteinase-28 proprotein; proenzyme form"
434..454
/note="Region: pro-sequence"
521..535
/note="Region: furin recognition sequence"
```

mat\_peptide 536..1729  
 /product="matrix metalloproteinase-28"  
 /note="active form"  
 misc\_feature 659..670  
 /note="putative N-glycosylation site"  
 misc\_feature 887..919  
 /note="Region: catalytic center"  
 misc\_feature 1232..1243  
 /note="putative N-glycosylation site"  
 3'UTR 1733..1817  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,39e-111 Length: 1817  
 Score: 2792.00 Matches: 519  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 98.52% Indels: 2  
 Gaps: 0

US-10-791-980-6 (1-520) x AF219624 (1-1817)

QY 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTProGlyHisLeu 20  
 Db 170 ATGTGTGCGCGCGCTGCGCTCTGCTGCGCGCGCTGCGAGCTGCTACTGTGGGCCCACTTG 229  
 QY 21 AppAlaGlnProAlaGlnArgGlyGlyGlnGlnLeuArgGlySerGlnAlaGlnAlaPheLeu 40  
 Db 220 GACGCCCGCGCGCGAGCGCGAGCGCGAGCTGGCGCAAGAGCGAGCGAGCGATTCCTTA 289  
 QY 41 GlnuYrYrGlyYrYrLeuAaenGlnGlnValProValaProThrSerThrArgPheSer 60  
 Db 290 GAGAAAGTACGAGATCTCTCAATGAACAGGTCCCAAGCTCCCACTCCCACTGATTGAC 349  
 QY 61 AspAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuAspArg 80  
 Db 350 GATGCCATCAGAGGGTTTCAGTGGGTGCCAGCTACCTTGACGGCGGTGGAGCGCG 409  
 QY 81 AlaThrLeuArgGlnMetThrArgProArgGlyValThrAspThrAsnSerTyrAla 100  
 Db 410 GCCACCTCGCGCAAGATGACTCTCCCGCTCGGGGTACAGATACCAACAGTTATGCG 469  
 QY 101 AlaTrpAlaGlnArgIleSerAspLeuPheAlaArgHisArgGlnTrpMetArgArgLys 120  
 Db 470 GCCTGGCTGAGAGATCATGACTTGTGTTGCTTAACACCGGACCAAAATGAGCGCTAAG 529  
 QY 121 LysArgPheAlaLysGlnGlyAsnLysTrpTyrIleGlnHisLeuSerTyrArgLeuVal 140  
 Db 530 AAACGCTTTCAAAGCAAGTAAACAAATGTACAAGCAGCACCTCTCTACCGCTGGTG 589  
 QY 141 AsnTrpProGlnHisLeu-ArgSerArgGlnPheGlyValaProCysAlaProProSerSe 160  
 Db 590 AACTGGCCTGAGCATCTGCCGAGCGCGCACTTGGGAGCGCGCGCGCGCTTCCAG 649  
 QY 160 rCysGlyAlaThrSerGlnArgTrpSerSerGlyValArgProGlnProGlnAlaProLeuTh 180  
 Db 650 TTGTGGAGCAACGTCACGCTGAGATCTGGGAGCGCGCCACCAAGCGCGCGCTGAC 709  
 QY 180 rSerGlySerProSerSerIleGlyThrThrMetGlyTyrTrpAlaMetProLeuMetAl 200  
 Db 710 ATCCGCTCACTCTTCCAAAGGAGCAACAAAGAGGCTGGGCAATGCCCTTATGAGC 769  
 QY 200 aGlnGlyAlaProTrpArgTrpProPheLeuProArgArgGlyGlnAlaHisPheAspGln 220  
 Db 770 CCAAGGGGCGCGCTGGGCGACCGCTTC-CTGCCCGCGCGCGCGCAAGCGCACTTCAACA 828  
 QY 220 nAspGlyArgTrpSerLeuSerArgArgArgIleArgAsnLeuPheValValLeuAlaHis 240  
 Db 829 AGATGAGCGCTGCTCTGAGCGCGCGCGCGCGCAACCTGTTCGTGCTGGCGCA 888  
 QY 240 sGlnIleGlyHisThrLeuGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaPr 260  
 Db 889 CGAGATGGTCAACGCTTGGCTCAACCACTCGCGCGCGCGCGCTCATGGCGCC 948

QY 260 cTyrTyrLysArgLeuGlyArgAspAlaLeuLeuSerTrpAspAspValLeuAlaValGln 280  
 Db 949 CTAATCAAGAGAGCTGGGCGCGCGAGCGGCTCTCACTGAGGAGCAACGCTGGCGCTGCA 1008  
 QY 280 nSerLeuTyrGlyLysProLeuGlnGlySerValaValGlnLeuProGlyLysLeuPh 300  
 Db 1009 GAGCTGTATGGAGAGAGCGCTTGAAGGCTCAAGTGGCGCTCCAGTCCCAAGAAAGCTGT 1068  
 QY 300 eThrAspPheGlnThrTrpAspSerTyrSerProGlnGlyArgArgProGlnThrGlnG 320  
 Db 1069 CACTGACTTGAAGACTGGGACTCTTACAGCGCCCAAGAGAGCGCTGAAACGAGGG 1128  
 QY 320 rProLysTyrCysHisSerSerPheAspAlaIleThrValaAspArgGlnGlnLeuTyr 340  
 Db 1129 CCTTAAATACGCAACTCTTCTGATGCGATCATCTGATGACAGGCAACAGCACTGTA 1188  
 QY 340 rIlePheLysGlySerHisPheTrpGlnValaAlaAlaAspGlyAsnValSerGlnProAr 360  
 Db 1189 CATTTTAAAGGAGCAATTTCTGAGAGTGGCGAGTGAATGGCAACGCTCAGAGCCCCG 1248  
 QY 360 rProLeuGlnGlyArgTrpValGlyLeuProProAsnIleGlnAlaAlaValaSerLe 380  
 Db 1249 TCACCTGCAAGAAATGGGTGGGCTGGCGCCCAACATTGAGCTGGCGAGCTCATTT 1308  
 QY 380 uAsnAspGlyAspPheTyrPhePheLysGlyGlyValArgCysTrpArgPheArgGlyProLy 400  
 Db 1309 GAATGATGAGATTTCTACTTCTTCAAGGGGCTGATGAGATTCGGGGCCCCCA 1368  
 QY 400 sProValTrpGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHisProAspAl 420  
 Db 1369 GCGAGTGGGGTCTTCCACAGCTGTGGCGGAGGGGGCTTGGCGCCGCACTGTAGCGC 1428  
 QY 420 aAlaLeuPhePheProProLeuArgArgLeuIleLeuPheLysGlyValaArgTyrVal 440  
 Db 1429 CGCCCTCTTCTCTCTCTGCGCGCGCTCATCTTCTTCAAGGGTCCCGCTTACAGT 1488  
 QY 440 lLeuAlaArgGlyGlyLeuGlnValaGluProTyrTyrProArgSerLeuGlnAspTrpG 460  
 Db 1489 GCTGCGCGAGAGGAGCTGCAAGTGGAGCCCTTACTACCCCGAAGCTGAGAGCATGGGG 1548  
 QY 460 rGlyIleProGlnGlnValSerGlyAlaLeuProArgProAspGlySerIleIlePhePh 480  
 Db 1549 AGGATCTCTGAGAGGTGAGCGCGCTGCGCGAGCGCGCGATGGCTCATCATCTTCTT 1608  
 QY 480 eArgAspAspArgTyrTrpArgLeuAspGlnAlaLysLeuGlnAlaThrThrSerGlyAr 500  
 Db 1609 CCGAGATGACCGCTACTGGCGCTTGCACCAAGCCAAATGACGACCACTCGGGCGC 1668  
 QY 500 gTrpAlaThrGlnLeuProTrpMetGlyCysTrpHisAlaAsnSerGlySerAlaLeuPh 520  
 Db 1669 CTGGGCGACCGAGCTCCTGATGGGCTGTGGCATGCAACCTCGGGAGCGCGCTGTT 1728  
 QY 520 e 520  
 Db 1729 C 1729

RESULT 6  
 AR528524  
 LOCUS AR528524 1985 bp DNA linear PAT 08-OCT-2004  
 DEFINITION Sequence 143 from patent US 6725730.  
 ACCESSION AR528524  
 VERSION AR528524.1 GI:53916602  
 KEYWORDS  
 SOURCE  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1985)  
 AUTHORS Bollinger,C.L., Jr.  
 TITLE Crane test weight assembly and method  
 JOURNAL Patent: US 6725730-A 143 27-Apr-2004;  
 FEATURES Location/Qualifiers  
 1..1985



ORIGIN /organism="unknown"  
/mol\_type="genomic DNA"

## Alignment Scores:

Pred. No.: 1,51e-111 Length: 1985  
Score: 2792.00 Matches: 519  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 98.52% Indels: 2  
DB: 6 Gaps: 0

US-10-791-980-6 (1-520) x AR528524 (1-1985)

Qy 1 MetValAlaArGValGlyLeuLeuLeuArGAlaLeuGlnLeuLeuLeuTrpGlyHisLeu 20  
Db ATGATCGCGCCGCTCGGCTCTGCTGCGCCGCTGCAAGCTGCTACTGTGGGGCCACTG 265  
Qy 21 AspAlaGlnProAlaGlnArgGlyGlyGlnGluLeuArgGlyGlyAlaGlnAlaPheLeu 40  
Db GACGCCCGACCGCGCGAGCGCGAGCGAGCTGCCCAAGAGCGCGAGGCAATTCCTA 325  
Qy 41 GlnuylTrpGlyTrpLeuAsnGlnGlnAlaProlyAlaProThrSerThrArgPheSer 60  
Db 326 GAGAAATACGAGATACCTCAATGAACAGGTCCCAAGCTCCCACTCCATCGATTACG 385  
Qy 61 AspAlaIleArGAlaPheGlnTrpValSerGlnLeuProValSerGlyValIleuAspArg 80  
Db 386 GATGCCATCAAGAGCTTTCAGTGGGTGTCACAGCTTCTGACGGCGGTGGACCGC 445  
Qy 81 AlaThrLeuArGlnMetThrArgProArgCysGlyValIleAspThrAsnSerTyrla 100  
Db 446 GCCACCGCGCCAGATGACTGCTCCCGCTGGGGGTACAGATACCAACAGTTATGCG 505  
Qy 101 AlaTrpAlaGlnuylSerAspLeuPheAlaArgHisArgThrIleuMetArgArglys 120  
Db 506 GCTGGGCTGAGAGATCAATGACTTGTTCCTACACACCGGACCAAAATAGAGCGTAA 565  
Qy 121 LysArgPheAlaLysGlnGlyAsnLysTrpTrpLysGlnHisLeuSerTyrlaArgLeuVal 140  
Db 566 AAACGCTTTCGAAAGCAAGGTACAAATGATGACAGACGACCTTCTTACCGCTGATG 625  
Qy 141 AsnTrpProGlnHisLeu-ArgSerArgGlnPheGlyAlaProCysAlaProProSerSe 160  
Db 626 AACTGGCGTGAACATCTGCCGAGCGCGAGTGGGGCGCGCTGCGCGCGCTTCAG 685  
Qy 160 TrpGlyAlaThrSerGlnArgTrpSerSerGlyValArgProGlnProGlnAlaProLeuTh 180  
Db 686 TTGTGGAGCAACGTCTCAGCGCTGAGTTCGGGAGGCCCAAGCCACAGGCCCGCTGAC 745  
Qy 180 TrSerGlySerProSerSerLysGlyTrpThrThrMetGlyTrpAlaMetProLeuMetAl 200  
Db 746 ATCCGCGTCACTTCTTCAGAGGGAGCAACAGATGGGCTGGGCAATGCCCTTGATGAC 805  
Qy 200 AGlnGlyAlaProTrpArgTrpProPheLeuProArgArgGlyGlyAlaHisAspAspGly 220  
Db 806 CCAAGGCGCGCGCTGGCGCACGCGCTTC-CTGCGCGCGCGCGGCAAGGCACTTGACCA 864  
Qy 220 AspGlnuylArgTrpSerLeuSerAspArgArgGlyValArgAsnLeuPheValIleuAlaHis 240  
Db 865 AGATGAGCGCTGGTCCCTGAGCGCGCGCGCGGCGCAACCTGTCGTGGTGGTGGCCCA 924  
Qy 240 AGlnIleGlyHisThrLeuGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaPro 260  
Db 925 CGAGATCGGTACACGCTTGGCTTCAACCACTCGCGCGCGCGCGCGCTCAAGGCC 984  
Qy 260 CTrpTrpTrpLysArgLeuGlyArgAspAlaLeuLeuSerTrpAspArgValIleuAlaValGly 280  
Db 985 CTACTACAAAGAGGCTGGCGCGGACGCGCTGCTGAGCGAGCGAGCTGCGCGCTGCA 1044  
Qy 280 nSerLeuTrpGlyLysPheLeuGlyGlySerValAlaValGlnLeuProGlyLysLeuPhe 300  
Db 1045 GAGCTGTATGGGAAGCCCTTAAGGGGCTCAGTGGCGGTCCAGCTCCAGGAAGCTGTT 1104

Qy 300 eThrAspPheGlnuylTrpAspSerTyrlSerProGlnGlyValArgArgProGlnuylGlnGly 320  
Db 1105 CACTGACTTTGAGACTGGAGCTCTTAACAGCCCGCAAGAGAGCGCCCTGAACGACAGG 1164  
Qy 320 TrpProLysTrpCysHisSerSerPheAspAlaIleThrValAspArgGlnGlnLeuTy 340  
Db 1165 CCTTAATACAGCCACTTCTCTCGATGCCATCATCTGTACACAGGCAACGCAACTGTGA 1224  
Qy 340 TrlPheLysGlySerHisPheTrpGlyValAlaAlaAspGlyAsnValSerGluProAr 360  
Db 1225 CATTTTAAAGGAGCCATTTCTGGAGGTGGACGCTGATGGCAACGTCTAGAGCCCCG 1284  
Qy 360 GProLeuGlnuyluylArgTrpValGlyLeuProProAsnIleGlnuylAlaAlaValSerLe 380  
Db 1285 TCACCTGAGAAAGAGAGGTGGGCTGGCCCCCAACATTAGGCTGGCGAGGTGCTATT 1344  
Qy 380 uAsnAspGlyAspPheTrpPhePheLysGlyValArgCysTrpArgPheArgGlyProly 400  
Db 1345 GAATGATGAGATTTCTTACTTTCCTCAAGGGGCTGATGCTGAGGTTCGGGGCCCCAA 1404  
Qy 400 sProValTrpGlyLeuProGlnLeuGlyArgAlaGlyGlyLeuProArgHisProAspAl 420  
Db 1405 GCCAGTGGGCTCCCAAGCTGTGCGGGCAAGGGGCTGCGCCGCTGACCTGACGC 1464  
Qy 420 AlaLeuPhePheProProLeuArGArgLeuIleLeuPheLysGlyAlaArgTyrla 440  
Db 1465 CGCCCTCTTCTTCCCTCTGCGCGCGCTCATCTCTTCAAGGGTGGCGCTACTAGT 1524  
Qy 440 IleuAlaArgGlyGlyLeuGlnuylGluProTyrlTrpProArgSerLeuGlnuylTrpGly 460  
Db 1525 GCTGGCGCGAGGGGACTGCAAGTGAAGCCCTTACTACCCCGAAGTGTGCGAGCTGGG 1584  
Qy 460 YGlyTrlEProGlnuyluylSerGlyAlaLeuProArgProAspGlySerTrlIlePhePhe 480  
Db 1585 AGGCATCTCGAGAGGTCAAGCGGCGCTGCCAGAGCGCGCATGCTTCATCATCTTCTT 1644  
Qy 480 eArgAspAspArgTyrlTrpArgLeuAspGlnuylAlaLysLeuGlnuylAlaThrThrSerGlyAr 500  
Db 1645 CCGAGATGACCGCTACTGGGCGCTCGACAGCGCAACCTGAGGCAACCACTCGGGCG 1704  
Qy 500 GTrpAlaThrGlnuyluylProTrpMetGlyCysTrpHisAlaAsnSerGlySerAlaLeuPhe 520  
Db 1705 CTGGGCGCACCGAGCTGCTGATGAGGCTGCTGCGCATGCCAACTCGGGAGCGCGCTGT 1764  
Qy 520 e 520  
Db 1765 C 1765

RESULT 7  
AX464010 1985 bp DNA linear PAT 16-JUN-2002  
LOCUS  
DEFINITION Sequence 143 from Patent WO0140466.  
ACCESSION AX464010  
VERSION AX464010.1 GI:21899022  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Baker,K.P., Bersini,M., Deforge,L., Desnoyers,L., Filvaroff,B.,  
Gao,W.Q., Gerlitsen,M.E., Goddard,A., Goddard,P.J., Gurney,A.L.,  
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,  
Wood,W.L. and Zhang,Z.  
Secreted and transmembrane polypeptides and nucleic acids encoding  
same  
Patent: WO 0140466-A 143 07-JUN-2001;  
JOURNAL Genentech Inc. (US)  
FEATURES  
Location/Qualifiers  
1..1985  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"

1.51e-211

|                        |         |               |      |
|------------------------|---------|---------------|------|
| score:                 | 2792.00 | meanln:       | 1985 |
| Percent similarity:    | 99.64%  | Matches:      | 519  |
| Best local Similarity: | 99.62%  | Conservative: | 0    |
| Query Match:           | 98.52%  | Mismatches:   | 1    |
| DB:                    |         | Indels:       | 2    |

AX464010 (1-1985)

|    |      |   |      |
|----|------|---|------|
| QY | 1    | MctValAlaAArgValGlyLeuLeuLeuLeuArgAlaLeuGlnLeuLeuLeuProGlnHisLeu  | 20   |
| Db | 206  | ATGCTCGCGCGCTGCGGCTCTCTCTCTCGCGCCCTCGAGCTGTACTGTGGGAGCCACCTG      | 26   |
| QY | 21   | AspAlaGlnProAlaGluArgGlyGlnGlnLeuLeuLysGlnAlaGlnAlaPheLeu         | 40   |
| Db | 266  | GACGCCCAAGCCCGCGGAGCCGCGAGGCCAGAGCTCGGCAAGAGCGGAGGACATTCTTA       | 326  |
| QY | 41   | GlnLysTyrGlyTyrLeuLeuGlnGlnValProLysAlaProThrSerThrArgPheSer      | 60   |
| Db | 326  | GAGAAAGTACGGATTCCTCATATGACAGAGTCCCAAGAGCTCCACCTCATCATTCAGC        | 385  |
| QY | 61   | AspAlaIleArgAlaPheGlnTyrValSerGlnLeuProValSerGlyValIleLeuArg      | 80   |
| Db | 386  | GATGCCATCAGAGCGTTTACAGGGGTATCCAGCTACTCTTCACGGCGGTGGACCCG          | 445  |
| QY | 81   | AlaThrLeuLeuArgGlnMetThrArgProArgCysGlyValThrAspThrAsnSerTyrAla   | 100  |
| Db | 446  | GCCACCCCTGCGCAATGACTGTCTCCCGCTCGGGGTACAGATACCAACATATTATGCG        | 505  |
| QY | 101  | AlaTyrPalAGluArgIleSerAspLeuPheAlaArgHisArgThrLysMetArgArgLys     | 120  |
| Db | 506  | GCCTGGGCTGAGAGGATCAGTGCATCTGTTGTCTAGACACCGGACCAAAATGAGGGGTAG      | 565  |
| QY | 121  | LysAspPheAlaLysGlnGlyAsnLysTyrTyrLysGlnHisIleLeuSerTyrArgLeuVal   | 140  |
| Db | 566  | AAACCCCTTGGCAAGAGCAAGGTAAACAAATGTTACAGACACACTCTCTCACGCTGTGTG      | 625  |
| QY | 141  | AsnTyrProGlnHisIleu-ArgSerArgGlnPheGlyValaProCysAlaProProSerSe    | 160  |
| Db | 626  | AACTGGGCTTAGCATCTGCGGAGCCGCGCATTTCCGGGGCGGTGGCGCGGCTTTCAG         | 685  |
| QY | 160  | TCysGlyValaThrSerGlnArgTyrPheSerSerGlyArgProGlnProGlnAlaProLeuThr | 180  |
| Db | 686  | TTTGAGGAGCAAGCTCTCACGCCCTGAGATTCTGGGAGGCCCCAGCACAGGCCCGCTGAC      | 745  |
| QY | 180  | IleArgLysSerProSerSerLysGlyThrThrThreGlyTyrTrpAlaMetProLeuMetAl   | 200  |
| Db | 746  | ATCCGCTCACTCTTCCMAAGGGAGCACAAACGATGGCGTGGCAATGCTTTGATGGC          | 805  |
| QY | 200  | ArgGlnGlyAlaProTyrArgThrProPheLeuProArgArgGlyGlnAlaHisIlePheAspG  | 220  |
| Db | 806  | CCAGGGGGGCGCTCGGCGCACGCCCTTC-CTGCCCGCGCGGGAAGCCCACTTGCAGCA        | 864  |
| QY | 220  | AspGlnLysArgTyrPheSerLeuSerArgArgGlyArgAsnLeuPheValIleLeuAlaHis   | 240  |
| Db | 865  | AGATGAGCGCTGCTCTCGAGCCGCCGCCCGGGGCCCAACCTGTTGTGTGCTCTGACCA        | 924  |
| QY | 240  | AGlnIleGlyHisThrLeuGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaPro     | 260  |
| Db | 925  | CGAATGTGGTACACGCTTGCGCTCACCCACTCGCCCGCGCGCGCTCATGTGGCGCC          | 984  |
| QY | 260  | CtyTyrTyrLysArgLeuGlyArgAspAlaLeuLeuSerTyrPheAspValIleuAlaValG    | 280  |
| Db | 985  | CTTCTACTAAGAGGCTGGGCGCGCACCGCGCTGCTCAGCTGGGACACACGTGTGCGCGTGA     | 1044 |
| QY | 280  | AsnLeuLysTyrGlyLysProLeuGlyGlySerValAlaValGlnLeuProGlyLysLeuPh    | 300  |
| Db | 1045 | GACCTGTATGGAGAGCCCTCTGAGGGGCTCATGTGCGCTCAGCTCCCAAGAAAGCTGTT       | 1104 |

|    |      |  |      |
|----|------|--|------|
| QY | 300  | eThrAspHeGluThrTTPAspSerTySerProGInG1AArgArgProGluThrGInG1       | 320  |
| Db | 1105 | CACTGACTTGTAGAACTCGACTCGACTCTACAGCCCCCAAGAAAGCGCTCGAAACGACGG     | 1165 |
| QY | 320  | yPProLySTyCySHIsSerSerPheAspAlaIleThrValAspArgGInGInLeuTy        | 340  |
| Db | 1165 | CCCTAAATACAGCCACTCTTCTTCAGTGCATCTACTGTAGACAGCAACACACTGT          | 1225 |
| QY | 340  | rIlePheTySGIySerHISpHeTrpGluValAlaAlaAspG1yAsnValSerGluProAr     | 360  |
| Db | 1225 | CATTTTAAAGGAGGCCATTTCTGGAGAGGTGCAGCTGATGGCAACGTCTCAGAGCCCCG      | 1284 |
| QY | 360  | gProLeuGInGInuArgTrpValG1yLeuProProAsnIleGluAlaAlaValAsrLe       | 380  |
| Db | 1285 | TCCACTCAGAAAGATGGGTGGGGCTGCCCCCAACATGAGGTGGCGAGTGTATT            | 1344 |
| QY | 380  | uAsnAspG1yAspPheTyTrpPhePheTySGIyG1yValrGySTrPArGrPheArgG1yProLy | 400  |
| Db | 1345 | GAATGATGAGATTTTCTACTTCTTCAAGAGGGGTGATGTCTGAGAGTTTCGGGGCCCCA      | 1404 |
| QY | 400  | sProValTTPG1yLeuProGInLeuCyAsrAlAG1yG1yLeuProArgHISProAspAl      | 420  |
| Db | 1405 | GCCAGTGTGGGTCTTCCACAGCTGTGCGGGAGGGGGCTGCCCGGCATCTTGACCC          | 1464 |
| QY | 420  | aAlaLeuPhePheProPoleuAlaGArgLeuIleLeuPheTySGIyAlaArgTyTrpVal     | 440  |
| Db | 1465 | CGCCCTCTTCTTCCCTCTCTGCGCCCTCACTCTCTTCAAGGGGCGCGCTACTAGT          | 1524 |
| QY | 440  | IleuAlaArgG1yG1yLeuGInValG1uProTyTyTrpProArgSerLeuGInAspTrng1    | 460  |
| Db | 1525 | GCTGGCCCCGAGGGGAGCTGCAGAGTGAGCCCTACTACCCCGAAGCTGCAGAGACTGGGG     | 1584 |
| QY | 460  | yG1yTlIeProGInGInuValSerG1yAlaLeuProArGrProAspG1ySerIleIlePhePh  | 480  |
| Db | 1585 | AGGCACTCCAGAGAGGTCAAGCGCGCTCGCGAGGCCCATGGCTCTCACTACTTCTT         | 1644 |
| QY | 480  | eArgAspAspArgTyTrpArgLeuAsnGInAlaIleLeuGInAlaThrTrpSerG1yAr      | 500  |
| Db | 1645 | CCGAGATGACCGCTACTGGCGCTGACCAAGGCCAATCTGACGCAACACTCGGGCCG         | 1704 |
| QY | 500  | gTrpAlaThrGluLeuProTrpMetG1yCyArTrpHISAlaAsnSerG1ySerAlaLeuPh    | 520  |
| Db | 1705 | CTGGGCAACGAGCTGCTCTGTAGGTGTGTGGCATGCCAACTCGGGAGGCCCTGT           | 1764 |
| QY | 520  | e 520  |      |
| Db | 1765 | C 1765   |      |

|            |   |         |      |        |                 |
|------------|---|---------|------|--------|-----------------|
| LOCUS      | AY358987  | 1985 bp | mRNA | linear | Pri 03-OCT-2003 |
| DEFINITION | Homo sapiens clone DNA151775 metalloproteinase (UNG1893) mRNA,  |         |      |        |                 |
| ACCESSION  | complete cds.   |         |      |        |                 |
| VERSION    | AY358987  |         |      |        |                 |
| KEYWORDS   | AY358987.1 GI:37183091  |         |      |        |                 |
| SOURCE     | FLI -CDNA.  |         |      |        |                 |
| ORGANISM   | Homo sapiens (human)  |         |      |        |                 |
|            | Homo sapiens  |         |      |        |                 |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;   |         |      |        |                 |
| REFERENCE  | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  |         |      |        |                 |
| AUTHORS    | Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Bush,J.,<br>Chen,J., Chow,B., Chun,C., Crowley,C., Currell,B., Deuel,B.,<br>Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,<br>Heldens,S., Huang,A., Kim,H.S., Klimoski,L., Liu,Y., Johnson,S.,<br>Lee,J., Lewis,L., Liao,D., Mak,M., Robbie,B., Sanchez,C.,<br>Schonfeld,J., Seehagiri,S., Simone,L., Singh,J., Smith,V.,<br>Stinson,J., Vagts,A., Vandlen,R., Wetanabe,C., Wileand,D., Woods,K.,<br>Xie,M.H., Yanesura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,<br>Goddard,A., Wood,W.I. and Godowski,P.<br>The secreted protein discovery initiative (SPDI), a large-scale<br>Effort to identify Novel Human Secreted and Transmembrane Proteins:<br>A Bioinformatics Assessment |         |      |        |                 |





DB 1594 C 1594

RESULT 10  
AF330002 2484 bp mRNA linear PRI 07-FEB-2001  
LOCUS Homo sapiens matrix metalloprotease MMP25 mRNA, complete cds.  
ACCESSION AF330002  
VERSION AF330002.1 GI:12698851  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
SOURCE

1. .2484

2 (bases 1 to 2484)  
Southan,C. and Hughes,S.A.  
Direct Submission  
Submitted (19-DEC-2000) Bioinformatics Target Discovery, SmithKline  
Beecham Pharmaceuticals, Third Avenue, Harlow, Essex CM32 5AW, UK  
location/Qualifiers

|    |     |   |     |
|----|-----|---|-----|
| Qy | 81  | AlaThrLeuArgG1MeThrArgProArgCysG1ValThrAspThrAsnSerTYrAla     | 100 |
| Db | 516 | GCCACCCTCCGGCAGATGACTGTCCTCCCGCTGGGGGTTTACAGATACCAACAGTTATGCG | 575 |
| Qy | 101 | AlaTrpAlaGluArg1LeSerAspLeuPheAlaArgH1aArgThrLybMeCArgGlyb    | 120 |
| Db | 576 | GCCTGGGCGTGAAGATCATGTACTGTTTGCTAGACACGGACCAAAATGAAGCGGTAAg    | 633 |
| Qy | 121 | LybArgPheAlaLysGlnGlyAsnLybTrpTYrLysGlnH1bLeuSerTYrArgLeuVal  | 140 |
| Db | 636 | AAACCTTTGCAAAGCAGAAAGTAACAATGATGACACACACACTCTCTCCCTCGGTG      | 695 |
| Qy | 141 | AsnTrpProGlnH1bLeuArgSerArgGlnPheGlyAlaProCysAlaProProSer     | 160 |
| Db | 696 | AACTGCGCTGACATCTCGCGAAGCGGAGTTCGGAGGCGCGCTGGCGCGCTTCAG        | 755 |
| Qy | 160 | rCysGlyAlaThrSerGlyArgTrpSerSerGlyArgProGlnProGlnAlaProLeuTh  | 180 |
| Db | 756 | TTGTGAGCAACGCTCTCAAGCTGTGAAGTTCTGGAGAGCCCCACACAGCCCCGCTGAC    | 815 |
| Qy | 180 | rSerGlySerProSerSerLysGlyThrThrThrMetGlyTrpAlaMeProLeuMetAl   | 200 |
| Db | 816 | ATCCGGCTCACTTCTTCCAAAGGGAGCCAAAGAGTGGGCTGGGCAATGCTTTGATGGC    | 875 |
| Qy | 200 | agInGlyAlaProTrpArgThrProPheLeuProArgArgGlyGluAlaH1bPheArgG1  | 220 |

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11"
CDS
276..1838

```

| Db | Seq  | Start | End | Score  |
|----|--|-------|-----|--------|
| Db | 876 CCAAGGGGGGCGCCCTG6CGCAGCGCTTC-CTGCCCCCGCGGGGGAAGCGCACTTCAGCA   | 93    | 141 | 100.00 |
| Qy | 220 nAaPGLuRgTTPSerLeuSerArGrArGrGjYarGaAnLeuPheValValLeuAaH       | 240   | 300 | 100.00 |
| Db | 935 AGATAGCGCGCTGTCTCCCTTAGCGCGCGCGCGCGGCGCAACTGTGTGTGTGTCTGCGCGCA | 99    | 147 | 100.00 |

```

/codon_start=1
/product="matrix metalloprotease MMP25"
/protein_id="AAK01706.1"
/db_xref="gi:12698852"
/translation="MVAVRGILLRLAQILLMGHLDQAQPAERCGQELKREKAAFLFKTG
YLNVEPRAPATSTPSDAIRAFQWVSQLYPQGVADLRQMTBPCGVDTVNSYAAIA
AERISDLFARHRTKMRKKRPAKQGNKYKQHSLSRLVNWBEHLPEAVRVAQVAAC
LMSVNSALEFWBAPATGPADIRLTFFQCDNHGSDNPDGCGALAAFLPRRBEAFL
DQDEWWSLSRRGRRLFVYLAHIEGHTGLTHSDPAFLMYPKALGRDALLSMWDVY
RQOOLYLKGPFGSGVAVQDQKGLFDFETDQWVSQSGGRPTQPGKCHSFEDITAY
RAGUUYLFGKSHFWVADQGVNSRPRLOEHWVCLPNIETAAVSLNDGDPFGKGR
CMRRGPKPVWNGLPQLCRAGGLPRHPDALEFFPLRLILFKGARYTYTLANGGLQVEF
YTPNSLDQWGIPEEVSGLPRPDGSIIFPDRIWRDLQAKLQATTSGRWATELPMWV
GCMANSGSALF"

```

|    |      |   |      |
|----|------|---|------|
| Oy | 240  | sglulleglyhsthrleuglyleuthrhisserploalaproargalaleumelap          | 260  |
| Db | 995  | cgaghtccgtcacacagcttgagctcacccaactgcgcggcgcgcttaccggcc            | 105  |
| Oy | 260  | ctyrrlylsargleuglyarqaspaalaleuleusertrpaspaspvalleualavalgl      | 280  |
| Db | 1055 | ctactacaaagagctgcggcgcgagccgcgtctgcactgcagcgaactgctgcgcgcga       | 1110 |
| Oy | 280  | nserleuylrglylvsproleuglvglservalalavalglleuproglylylsleuph       | 300  |
| Db | 1115 | gagcctgthnrggaagagcccttgcggggcctcagtgccgcctccacactccacagaaagctgtt | 1170 |
| Oy | 300  | ethrasppegelunhrttrpaspserlyserpprognglyarqargproglnuthrgngl      | 320  |

|                        |           |
|------------------------|-----------|
| ORIGIN                 |           |
| Alignment Scores:      |           |
| Pred. No.:             | 1,876-111 |
| Score:                 | 2792.00   |
| Percent Similarity:    | 59.62%    |
| Best Local Similarity: | 59.62%    |
| Query Match:           | 58.52%    |
| OB:                    | 9         |
|                        |           |
| Length:                | 2484      |
| Matches:               | 519       |
| Conservative:          | 0         |
| Mismatches:            | 1         |
| Indels:                | 2         |
| Gaps:                  | 0         |

Db 1175 CACTGACTTTGAGACCTGGGACTCCCTACAGCCCCCAGGAAGGCGCCTCGAAACGACAGG 1230

Qy 320 YProLysrYrCySHiSseSerPheAspAlaIleThrValaIapArgGInGInLeuTy 340

Db 1235 CCTTAAATACCTCCACTCTTCTTGATGCCCATCTGTAGACGGCAGACGAACTGTA 1290

Qy 340 rIlePheLysGjYserHilapHeTIpGluValAlaalaapGjYasnValSerGluProAr 360

US-10-791-980-6 (1-520) X AF330002 (1-2484)

Db 1295 CATTTTTAAAGGAGCCATTCTGGGAGGTGGCAGCTGATGGCAACGTCTCAGAGCCCG 135

1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHisIleu 20

360 gptleugncinwagltipalglleupropfoamileginalaiaaavalseide 380

Db 276 ATGTGCGCGCGGCTCGGCTCTGTGCGCGGCCCTGCAGTGTCTACTGTGGGCCACCTG 335

[illegible]

21 AspaLaGInProLaGInArgGlyGInGluLeuArgGlyGInLaGInLaPheLeu 40

Db 1415 GANTGAGAGATTCTTCTTCAAAAGGGGTCGATGCTGAGGTTCCGGGGCCCCAA 147

Db 336 GACGCCAGCCCGCCGAGCGCCAGAGCCCAAGACTGCGCAAGAGGCGGAGGCATTCCCTA 395

QY 400 sProlValTyrGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHisProAla 420

41 Glycyltyrosyltyrosylleuanylglutylvalylprolylvalylprothylserthirgphenyl 60

Db 1475 GCGAGTGTGGGGTCTCCACAGCTGTGCCGGCAGGGGGCCTGTCCCCGCCATCCTGACGC 153

[illegible]

420 aaAlaLeuPheProPLeuArgArgLeuIleuPheGlyAlaArgTyrTyrVa 440

456 GATGTCATAGAGCGTTTCAGTGGATGCCAGCTACCTCTACGGGCGCTGTGAACCGC 515

Db 1535 CGCCCTTCTTCCCTCCTCGCGCGCCATCCTTCAAGGGTCCCGCTACTAGT 159

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

QY 440 IleuAlaArgGlyGlyLeuGlnValGluProTyrTyrProArgSerLeuGlnAspTrpGlu 460

Db 1595 GCTGCCCGAGGAGGAGCTGCAAGTGAAGCCCTTACTACCTCCGGAAGCTCCAGAGCTGGGG  
 Qy 460 YGlyIleProGluGluValAlaLeuProArgProAspGlySerIleIlePhePh 1654  
 Db 1655 AGGCATCCCTGAGAGGAGTCAAGCGGCGCTCCGAGGCGCGAGTGGCTCATATCTTT  
 Qy 480 eArgAspAspArgTyrTrpArgLeuAspGlnAlaIleuGlnAlaIleuTrpSerGlyAr 1714  
 Db 1715 CCGAGATGATCCGCTACTGCGCTCGACAGGCGCAAACTGCGAGCAACCACTCGGGCGG 500  
 Qy 500 GTTAlaThrGluLeuProTrpMetGlyCysTrpHisAlaIleuSerGlySerIleIlePhePh 1774  
 Db 1775 CTGGGCCACCGAGCTGCGCTCGATGGGCTGTGTCATGCAACTCGGGGAGCGGCTTT 520  
 Qy 520 e 520  
 Db 1835 C 1835

RESULT 11  
 AX398365 2527 bp DNA linear PAT 27-MAY-2002  
 LOCUS AX398365  
 DEFINITION Sequence 1 from Patent WO0220739.  
 ACCESSION AX398365  
 VERSION AX398365.1 GI:21261133  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM

REFERENCE  
 AUTHORS Human matrix metalloproteinase  
 TITLE Patient: WO 0220739-A 1 14-MAR-2002;  
 JOURNAL Millennium Pharmaceuticals, Inc. (US)  
 FEATURES  
 source  
 1. 2527  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.9e-111 Length: 2527  
 Score: 2792.00 Matches: 519  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 98.52% Indels: 2  
 Gaps: 0

US-10-791-980-6 (1-520) x AX398365 (1-2527)

Qy 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyIleu 20  
 Db 300 ATGGTGGCGCGCGTGGCGCTCTGCTGGCGCCCTCGACGCTGCTACTGTGGGCCACTG 359  
 Qy 21 AspAlaGlnProAlaGluArgGlyGlyGlnGluLeuArgGlyGlnAlaGlnAlaPheLeu 40  
 Db 360 GAGGCCAGCCCGCGAGAGCGCGAGGCAAGAGCTGCGCAAGAGGCGGAGCAATTCCTA 419  
 Qy 41 GlnGlyArgGlyTyrLeuAsnGlnGlnValProIleuAlaProThrSerThrArgPheSer 60  
 Db 420 GAGAAAGTACGATACCTCAATGAAAGAGTCCCAAAAGTCCCACTCCACTCGANTCAGC 60  
 Qy 61 AspAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuAspArg 80  
 Db 480 GATGCAATCAGAGCGTTCAGTGGGTGTCCAGCTACCTGTCAAGCGGCGTGTGGACCGC 539  
 Qy 81 AlaThrLeuArgGlnMetThrArgProArgCysGlyValThrAspThrAsnSerTyrAla 100  
 Db 540 GCCACCTTGCGCAGATGACTCGTCCCGCTCGGGGTTACAGATACCAAGTTATGGG 599  
 Qy 101 AlaTrpAlaGluArgIleSerAspLeuPheAlaArgHisArgThrIleuMetArgArgIys 120

Db 600 GCTTGGGCTGAGAGATCACTGACTGTGTTGCTGACACCGGACCAAAATGAGGCGTAA 659  
 Qy 121 LysArgPheAlaIleuGlnGlyValLeuIleuTyrIleuGlnIleuSerTyrArgLeuVal 140  
 Db 660 AAACGCTTTCGAAAGCAAGGTTAAACAAATGATGTAACAAGCAGACCTCTCTACCGCTGG 719  
 Qy 141 AsnTrpProGlnIleuArgSerArgGlnPheGlyValProCysAlaProPheSer 160  
 Db 720 AACTGGCTGAGCATCTGCGAGCGGACAGTGGGGCGCGGCGCGGCTTCCAG 779  
 Qy 160 TCGGlyAlaIleuSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuPh 180  
 Db 780 TTGTGAGCAACGCTCAAGCGCTGAGATTTGAGGAGGCCCAAGCCCAAGGCCCGCTGAC 839  
 Qy 180 rSerGlySerProSerSerIleuGlyIleuTrpMetGlyTyrAlaMetProLeuMetAl 200  
 Db 840 ATCCGGCTCACTCTCTTCCAAAGGGAGCAACAGATGGGCTGGGCAATGCTTTGATGCG 899  
 Qy 200 aglnGlyAlaProTrpArgTrpPheLeuProArgArgGlyGlyIleuAlaIleuAspGly 220  
 Db 900 CAGGGGCGCGCTGGCGCAGCGCTTC-CTGCCCGCGCGCGCAAGCGCACTTCGACCA 958  
 Qy 220 nAspGluArgTrpSerIleuSerArgArgArgGlyArgAsnLeuPheValIleuAlaIle 240  
 Db 959 AGATGAGCGCTGGTCTGAGCGCGCGCGCGCAACCTGTTCTGCTGCTGGCGCA 1018  
 Qy 240 sGlnIleuGlnIleuTrpLeuGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaPr 260  
 Db 1019 CGAGATCGGTCAACGCTTGGCTTCACCACTGCGCGCGCGCGCGCGCTCTATGCGCGC 1078  
 Qy 260 cTyrTyrIleuArgGlyArgArgAlaLeuLeuSerTrpAspAspValLeuAlaIle 280  
 Db 1079 CTACTCAAGAGGCTGGCGCGCGCGCTGCTGAGTGGAGCAAGCTGCTGGCGCTGCA 1138  
 Qy 280 nSerLeuTyrGlyLeuProLeuGlyGlySerValAlaValGlnLeuProGlyIleuLeuPh 300  
 Db 1139 GAGCTGTATGGAGAGCCCTAGAGGCGCTCAGCGCTCAGAGCTCCAGGAAAGCTGTT 1198  
 Qy 300 eThrAspPheGlnIleuTrpAspSerTyrSerProGlnGlyArgArgProGlnTrpGln 320  
 Db 1199 CACTGACTTGAAGCTGGGACTCTTACAGCCCCCAAGAGGCGCGCTCAAGAGCGCGC 1378  
 Qy 320 YProIleuTyrCysHisSerSerPheAspAlaIleuTrpValAspArgGlnGlnLeuTyr 340  
 Db 1259 CCTAAATACTGCACTCTCTTGCATGCCATCACTGTAGACAGCAACCACTGTA 1318  
 Qy 340 rIlePheIleuGlySerHisPheTrpGlnValAlaAlaAspGlyAsnValSerGlnProAr 360  
 Db 1319 CATTTTAAAGGAGCAATTTCTGGAGGTGGAGCTGATGGCAAGCTCAGAGCCCGC 1378  
 Qy 360 gProLeuGlnGluArgTrpValGlyLeuProProAsnIleuAlaAlaIleuValSerIle 380  
 Db 1379 TCCACTGCAAGAAAGATGGTCCGGCTGCCCAATTTGAGGCTGGCGGAGTCTATT 1438  
 Qy 380 uAsnAspGlyAspPheTyrPhePheIleuGlyGlyArgCysTrpAspPheArgGlyProIy 400  
 Db 1439 GAATGATGAAATTTCTACTTCTTCAAGGGGCTGATGCTGAGAGTTCCGAGGCCCA 1498  
 Qy 400 sProValTrpGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHisProAspAl 420  
 Db 1499 GCCAGTGTGGGTCTCCACAGCTGTCCGGCGAGAGGGGCGTGGCCCGCATCTGACGC 1558  
 Qy 420 aaIleuPhePheProProLeuArgArgLeuIleuPheIleuGlyValAlaArgTyrTyrVa 440  
 Db 1559 CGCCCTCTTCTCTCTCTGCTGCGCGCTCATCTCTTCAAGGGTCCCGCTACTACGT 1618  
 Qy 440 IleuAlaArgGlyGlyIleuGlnValGluProTyrTyrProArgSerLeuGlnAspTrpG 1618  
 Db 1619 GCTGGCCGAGGAGGAGCTGCAAGTGAAGCCCTACTACCCCGAAGCTCCAGAGCTGGGG 1678  
 Qy 460 YGlyIleProGluGluValSerGlyAlaLeuProArgProAspGlySerIleIlePhePh 480

|                        |   |  |                 |
|------------------------|---|--|-----------------|
| Db                     | 1679  | AGGATCCCTAGAGAGGTCACGGGGCCCTGCCAGAGGCCATGGCTCATCATCTTCTT       | 1738            |
| Qy                     | 480   | eaYgaPaPaPaYrYrTTPaTgLeuAaPcIaIaLyLeuGlaIaThrThirSerGlyAr      | 500             |
| Db                     | 1739  | CCGAGATACCGCTACTGCGCGCTCGAACGAGCCAAACTGGCAGGCAACACCTCGGCGC     | 1798            |
| Qy                     | 500   | gTTPaIaThrGtLeuPcPcTTPMeGtLyCySerThPthIaIaAaSerGtLySerAlaLeuPh | 520             |
| Db                     | 1799  | CTGGGCGCCACGAGCTGCGCTGGATGGCTGCTGGATCCCAACTCGGAGAGCGCTGTT      | 1858            |
| Qy                     | 520   | e 520  |                 |
| Db                     | 1859  | C 1859   |                 |
| RESULT 12              |   |  |                 |
| AF315683               |   | 1753 bp  | mRNA            |
| LOCUS                  |   |  | linear          |
| DEFINITION             | Homo sapiens matrix metalloproteinase 28 (MMP28) mRNA, complete   |  | PRI 06-APR-2001 |
| ACCESSION              | AF315683  |  |                 |
| VERSION                | AF315683.1  |  | GI:11935131     |
| KEYWORDS               |   |  |                 |
| SOURCE                 |   |  |                 |
| ORGANISM               | Homo sapiens (human)  |  |                 |
| REFERENCE              | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |  |                 |
| AUTHORS                | 1 (bases 1 to 1753)   |  |                 |
| TITLE                  | Marchenko, G.N. and Strongin, A.Y.  |  |                 |
| JOURNAL                | MMP-28, a new human matrix metalloproteinase with an unusual  |  |                 |
| MEDLINE                | cysteine-switch sequence is widely expressed in tumors  |  |                 |
| PUBMED                 | Gene 285 (1-2), 87-93 (2001)  |  |                 |
| REFERENCE              | 2 (bases 1 to 1753)   |  |                 |
| AUTHORS                | Marchenko, G.N. and Strongin, A.Y.  |  |                 |
| TITLE                  | Direct Submission   |  |                 |
| JOURNAL                | Submitted (23-Oct-2000) The Burnham Institute, 10901, North Torrey Pines Road, CA 92037, USA                                |  |                 |
| FEATURES               |   |  |                 |
| Source                 | Location/Qualifiers   |  |                 |
|                        | 1..1753   |  |                 |
|                        | /organism="Homo sapiens"  |  |                 |
|                        | /mol_type="mRNA"  |  |                 |
|                        | /db_xref="taxon:9606"   |  |                 |
|                        | /chromosome="17"  |  |                 |
|                        | /map="17q21.1"  |  |                 |
|                        | 1..1753   |  |                 |
|                        | /gene="MMP28"   |  |                 |
|                        | 16..1578  |  |                 |
|                        | /gene="MMP28"   |  |                 |
|                        | /note="MMP-28"  |  |                 |
|                        | /codon_start=1  |  |                 |
|                        | /product="matrix metalloproteinase 28"  |  |                 |
|                        | /protein_id="AA041981.1"  |  |                 |
|                        | /db_xref="GI:11935132"  |  |                 |
|                        | /translation="MWAVGGLLRALQLLWGHLDADPAERGGQELREKAAFLKRYG   |  |                 |
|                        | /feature_1="AA041981.1"   |  |                 |
|                        | YLRNPVAPASTRFSDAIRAFQWVSQLPVSGVLDRAITLQMTSPRCGVDTNSYAAAF  |  |                 |
|                        | AEISDLPABRTKRRKKRPKQGNKKYKQHLSTRLVWPEHLPEPVRGAVRAAF   |  |                 |
|                        | LMSVLSALEFWAPATGPGADIRITLPGDGHNDGNAFDGGALAHAF.PRRGEAHL  |  |                 |
|                        | DODKWSLSRRRGRNLFYVLAHEIHTTGLTHSPAPRLAMPYIKRLRDALLSHDDV  |  |                 |
|                        | LAVSLYKPKLGGSVAVQLPGCLFTDFETWDSYSPQGRPEYQKCHSSFDALTVN   |  |                 |
|                        | RQOQLYFKSGHFWADVNSBEPPLQERWVG.LPPIIEAAVSLNDGDFYFPGGGRV  |  |                 |
|                        | CMRFPGKPVWGLPOLCRAGGLPRPDALFPPPLRLILFGARVYVILAGGQLQVE   |  |                 |
|                        | YPRSLDMGGIPREVSQALPRPDSSILFPDRDRWCLDQKLQATTSGRNATELPMW  |  |                 |
|                        | GCMAHSSSALF"  |  |                 |
| ORIGIN                 |   |  |                 |
| Alignment Scores:      |   |  |                 |
| Pred. No.:             | 2,95e-111   | Length:  | 1753            |
| Score:                 | 2784.00   | Matches:   | 518             |
| Percent Similarity:    | 99.42%  | Conservative:  | 0               |
| Best Local Similarity: | 99.42%  | Mismatches:  | 2               |
| Query Match:           | 98.24%  | Indels:  | 2               |
| DB:                    | 9   | Gaps:  | 0               |

|    | US-10-791-980-6 (1-520) x AF315683 (1-1753)                               |
|----|---|
| QY | 1 MecValaIaargValaGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyHisLeu 20        |
| Db | 16 ATGGTCGGCGCGGTGGCGCTCTCGTCGGCGCGCTCGACAGTGTCTAGTGGGGCCACTGG 75         |
| QY | 21 AbspAaGlnProAlaGlnLysGlyGlyGlnGlnLysLeuAArgGlyGlnAlaGlnAlaPheLeu 40    |
| Db | 76 GACGCCCAAGCCCGGAGGCGGAGGCGGAGGCTGGCGGAGGAGGCGGAGGCAATTCCTTA 135        |
| QY | 41 GlnuysTrgLyTyTyLeuAaGlnGlnValaProLyAAlaProThSerThArgPheSer 60          |
| Db | 136 GAGAAAGTACGGATTACTCAATGAACAGGTCGCCAAGAGCTCCACCTCCACTGATTCACG 195      |
| QY | 61 AbspAlaIleAargAlaPheGlnTrpValSerGlnLeuProValSerGlyValaLeuAaParg 80     |
| Db | 196 GATGCGATCAAGAGCGCTTTCAGTGGGTGTCCGACTACCTGTACGGCGCTGTGGACCGC 255       |
| QY | 81 AlaThrLeuAArgGlnMetThrArgProArgCysGlyValaThrAaPThrAaMetTyTyAla 100     |
| Db | 256 GCGACCCCTGGCGCAATATCTGTCCCGCTGCGGGGTTCAGATATCCAAGATTATGGG 315         |
| QY | 101 AlaTrpAlaGlnArgIleSerAaPheAaPheAlaArgHisArgThrLysMetArgArgLys 120     |
| Db | 316 GCGTGGCGGTAGAGGAGATCAGTACCTTTGTTGTATGACACCGGACCAATATGAGCGTAA 375      |
| QY | 121 LysAaArgPheAlaLysGlnGlyAaAsnLysTrpTyTyGlnHisLeuSerTyArgLysLeuAl 140   |
| Db | 376 AAACCCCTTTCGAAAGCAAGGTAAACAAATGGTACAAAGCAAGCACTCTCTACCGCTGTG 435      |
| QY | 141 AaNTPrProGlnHisLeu - ArgSerArgGlnPheGlnAlaProCysAlaProProSerSe 160    |
| Db | 436 AACTGACCTTGACATCTGCGGAGCGCGGACATTCGGGGGCGCGTGGCGCGCTTCCAG 495         |
| QY | 160 rCysGlyAlaThrSerGlnArgTrpSerSerGlyValArgProGlnProGlnAlaProLeuTh 180   |
| Db | 496 TTGTGAGCAAGCTCTCAGCGCTGAGTTCTTGAGAGGCCCGGACCAAGCGCCCGCTGAC 555        |
| QY | 180 rSerGlySerProSerSerLysGlyThrThrMetGlyTrpAlaMetProLeuMetAl 200         |
| Db | 556 ATCCGCGCTCACCTTTCTTCGAAAGGAGCAACAAAGATGGCTGGGCAATCCCTTTGATGCG 615     |
| QY | 200 aGlnGlyAlaProTrpArgThrProPheLeuProArgArgGlyGlnAlaHisPheAaPgl 220      |
| Db | 616 CCAAGGGGCGCGCTTGGCGGACGCTTC - CTGCGCGCGCGCGCGGCAAGGCCACTTGCACA 674    |
| QY | 220 naSPGlnuArgTrpSerLeuSerArgArgArgGlyValArgAaLeuPheValValLeuAlaHis 240  |
| Db | 675 AGATGAGCGCGTGTCCCTTGAGCGCGCGCGCGGCAACCTGTTGATGTCTGCGCGCA 734          |
| QY | 240 aGlnLleGlnHisThrLeuGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaPr 260      |
| Db | 735 CGAGATCGGTCAACAGCTTGGCTTACCCACTCGCGCGCGCGCGCGCTCAATGGCGCGC 794        |
| QY | 260 cTyTrpTyTyAaArgLeuGlyArgAaPAlaLeuLeuSerTrpAaPArgValLeuAlaValGln 280   |
| Db | 795 CTACTATCAAGAGGCTGGGCGCGGACGCGCTGTCAAGCTGGACAGAGCTGTGGCGGTGA 854       |
| QY | 280 nSerLeuTyTrgLyLysProLeuGlyGlySerValaAlaValGlnLeuProGlyLysLeuPh 300    |
| Db | 855 GAGCTGTATGGGAAGCCCTTATGGGGGCTCAATGGCGCGTCCAGCTCCGAGAAAGCTGTT 914      |
| QY | 300 eThrAaPheGlnuTrpTrpAaPSerTySerProGlnGlyAaArgArgProGlnuThrGlnGln 320   |
| Db | 915 CACTACTTTGAGACTTGGACTGTCTTCAACAGCCCGCAAGAAAGCGCGCTGAAACGAGGG 974      |
| QY | 320 yProLysTyTyCysHisSerSerPheAaPAlaIleThrValaAaPArgGlnGlnGlnLeuTy 340    |
| Db | 975 CCTTAATATGTCCTCACTCTTCTTGTGATGTCATCACTGTAGAACAAGCAACAGAACTGTA 103     |
| QY | 340 rTlLePheLysGlySerHisPheTrpGlnuValaAlaAlaAaPArgLysAaValSerGlyProAr 360 |



Db 1035 CATTTTAAAGGAGCCATTCTGGAGGTGGCAGCTGATGCAACGCTCAGAGCCCG 1094  
 Qy 360 gProLeuGlnGluArgTTrpValGlyLeuProProAsn116GluAlaAlaValSerLe 380  
 Db 1095 TCCACTGTCAGAAAGATGGGTGGGCTGCCCCCAACATTGAGGCTGGCGAGTGCATT 1154  
 Qy 380 uAenAPGlyAaPhePheTyrPhePheGlyGlyValArgCytTrpArgPheArgGlyProLy 400  
 Db 1155 GAATGATGAGATTTCTACTCTTCAAAAGGGGGGTCAGATGTGAGGCTTCGGGGCCCCCA 1214  
 Qy 400 sProVal1TrpGlyLeuProGlnLeuArgAlaGlyGlyLeuProArgHisProAspAl 420  
 Db 1215 GCGAGTGTGGGATCTCCCAACGCTGCGGGGCGAGGGGCGCTGCCCCCATCTGAGCGC 1274  
 Qy 420 aAlaLeuPhePheProProLeuArgArgLeuLeuPhePheGlyAlaArgTyrTrpVal 440  
 Db 1275 CGCCCTCTTCTCCCTCTCTCTGCGCGCTCATCTCTTCAAGGGTCCCGCTACTACCGT 1334  
 Qy 440 lLeuAlaArgGlyGlyLeuGlnValGlyLeuProTyrTrpArgSerLeuGlnAspTrpG 460  
 Db 1335 GCTGGCCCGAGGGGGGAGCTGCAAGTGAAGCCCTACTACCCCGAAGTCTGAGAGCTGGGG 1394  
 Qy 460 yGly11LeuProGlnGluValSerGlyAlaLeuProArgProArgPlySerLeuLeuPheP 480  
 Db 1395 AGGCATCCCTGAGAGAGTCAAGGGGGGCTGCGAGGCGCGAGCTGCTCATCTCTCTT 1454  
 Qy 480 eArgAPaPAPArgTyrTrpArgLeuAPGlnAlaValSerGlnAlaAlaTrpTrpSerGly 500  
 Db 1455 CCGAGATACCGGCTACTGAGTGTGCTTCCGACGAGGCAACATGAGGAGACACTCTGGGGCG 1514  
 Qy 500 gTrpAlaTrpGluLeuProTrpMetGlyCysTrpTrpAlaAlaAsnSerGlySerAlaLeuP 520  
 Db 1515 CTGGGCGACCGAGCTGCTGCTGATGGGCTGCTGCGATCGCAACTGAGGAGACACTCTGGGGCG 1574  
 Qy 520 e 520  
 Db 1575 C 1575

RESULT 13  
 BC002631  
 LOCUS Homo sapiens matrix metalloproteinase 28, transcript variant 2, 2091 bp, mRNA, linear, PRI 29-JUN-2004  
 DEFINITION BC002631 mRNA (cDNA clone MGC:4164 IMAGE:3610296), complete cds.  
 VERSION BC002631.2 GI:33876957  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Hopkins, R.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., Diachenko, L., Maruina, K., Moore, T., Max, S.I., Wang, J., Hsieh, F., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tothlyuk, S., Abramson, R.D., Mullany, S.J., Bosak, S.A., Loguigliano, N.A., Peters, G.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., McEwan, P.J., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Richards, S., Villalón, D.K., Muzny, D.M., Sodergren, E., Lu, X., Gibbs, R.A., Sanchéz, A., Whitting, E., Kettner, M., Madan, A., Young, A.C., Rodriguez, S., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalski, U., Smalins, D.E., human and mouse cDNA sequences  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 PUBMED 12477932

## REFERENCE

2 (bases 1 to 2091)  
 Strausberg, R.  
 Direct Submission  
 Submitted (05-FEB-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 19, 2003 this sequence version replaced gi:12803592.  
 Contact: MGC help desk  
 Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Sequencing Center (NISC), National Institutes of Health Intramural  
 Gaithersburg, Maryland.  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-U., Karlins, E., Kwong, P., Latig, P., Legaspi, R.,  
 Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McElroy, J.C.,  
 McDowell, J., Pearson, R., Stanciro, S., Thomas, P.J., Touchman, J.W.,  
 Turgeson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

## REMARK

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/HLN at: <http://image.llnl.gov>  
 Series: IPAL Plate: 12 Row: a Column: 19  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, GenomScan gene prediction.  
 Location/Qualifiers  
 1..2091  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:4164 IMAGE:3610296"  
 /issue\_type="pancreas, adenocarcinoma"  
 /clone\_id="NM\_0339"  
 /lab\_host="DH10B-R"  
 /note="vector: pOTB7"  
 1..2091  
 /gene="MMP28"  
 /note="synonyms: MMP25, MMP28"  
 /db\_xref="LocusID:79148"  
 /db\_xref="MIM:608417"  
 248..1429  
 /gene="MMP28"  
 /codon\_start=1  
 /product="matrix metalloproteinase 28, preproprotein  
 /protein\_id="AAH02631.1"  
 /db\_xref="GI:12803593"  
 /db\_xref="LocusID:79148"  
 /db\_xref="MIM:608417"  
 /translation="MVARVGLLRALQLLIMGLDAPARGGQELRKEARAFLEKRG  
 AERISDLFAERTRMRRKRFAPQAKNRKQKHLSTRLVWVPHPIPERVAVGAFAAF  
 LMSVSLFAERWEPWAPGADIRLTFQGDHNGLGNAFDGGCLAHAFLEPRGEAF  
 DQDRMSLSRRRRNIFVLAHEIHTGLTSPAPRALMPYTKRIGRALSLMDVY  
 LAVQSLGKRLGSAVAVLPGLFTFDFTWDSYSGRPFRTQPKYCHSSPFAITVD  
 RQOGLYIFKSHPEVAADGNVSEBRPIQERNVGIPLNTEAAVASLNDGDFYFKVQS  
 V"

## FEATURES

## source

## gene

## CDS

## ORIGIN

## Alignment Scores:

Pred. No.: 1.2e-110  
 Score: 2771.50  
 Percent Similarity: 95.76%  
 Best Local Similarity: 95.76%

Length: 2091  
 Matches: 519  
 Conservative: 0  
 Mismatches: 1



Query Match: 97.79% Indels: 23  
 DB: 9 Gaps: 1  
 US-10-791-980-6 (1-520) x BC002631 (1-2091)

QY 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuLeuTrpGlyVHleu 20  
 DB 248 ATGTGCGCGCGCTGCGCTCTCTGCTGCGCGCTGCGAGCTGTACTGTGTGGGCGACCTG 307

QY 21 AspAlaGlnProAlaGluArgGlyGlyGlnGluLeuArgGlyValGlnAlaPheLeu 40  
 DB 308 GACCCCGAGCGCGCGCGAGCGAGCGAGGAGCTGCGCAAGAGAGCGAGCGACTTCTTA 367

QY 41 GluValTrpGlyTrpLeuAsnGlnGlnValProIysAlaProThrSerTrpArgPheSer 60  
 DB 368 GAGAGATACGATACCTCAATGAAACAGGTCCCAAGCTCCCACTTCCACTCGATTACG 427

QY 61 AspAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuAspArg 80  
 DB 428 GATCCCATCAGAGCGTTTCACTGGGTGTCCAGCTACTCTGTCAAGCGGTGTGACCGC 487

QY 81 AlaThrLeuArgGlnMetThrArgProArgCysGlyValThrAspThrAsnSerTrpAla 100  
 DB 488 GCCACCTCGCGCGAGATGACTGTCCCGCTGCGGAGTTACAGATACCAACAGTTATGCG 547

QY 101 AlaTrpAlaGluArgGlySerAspLeuPheAlaArgHisArgThrIysMetArgArgIys 120  
 DB 548 GCGTGGGCTGAGAGATGACTGTCTTCTCTAGACACCGGACCAAAATGAGCGCTAG 607

QY 121 LysArgPheAlaLysGlnGlyAsnLysTrpTrpLysGlnHisLeuSerTrpArgLeuVal 140  
 DB 608 AAAGCTTTGCAAGCAAGGTAAACAAATGTATCAAGCAGCACTCTCTACCGCTGGTG 667

QY 141 AsnTrpProGlnHisLeu-ArgSerArgGlnPheGlyValaProCysAlaProProSerSe 160  
 DB 668 AACTGGCGCTGACATCTGCGAGCGGCACTGCGGCGCGCGCGCGCTTCAG 727

QY 160 rCysGlyValaThrSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuTh 180  
 DB 728 TTGTGGAGCAACGTCTCAGCGCTGGAAGTTCTGGAGCGCCCAAGCAAGCGCGCTGAC 787

QY 180 rSerGlySerProSerSerLysGlyThrThrThrMetGlyTrpAlaMetProLeuMetAl 200  
 DB 788 ATCCGCGCTACCTTCTTCAAGGAGCAACAGATGGGCTGGGCAATGCTTGTATGCGC 847

QY 200 aGlnGlyAlaLeuProTrpArgThrProPheLeuProArgArgGlyGlnAlaHisPheAspG 220  
 DB 848 CCAGGGGCGCGCTGCGCGAGCGCTTC-CTGCCCCGCGCGCGAGAGCACTTCGACCA 906

QY 220 nAspGlyLysTrpSerLeuSerArgArgGlyArgGlnLeuPheValValLeuAlaHis 240  
 DB 907 AGATGAGCGCTGTCTCTGAGCGCGCGCGCGCGCGCACTTCTGTGTGTGTGCGCA 966

QY 240 aGluLysGlyHisThrLeuGlnGlyLeuThrHisSerProAlaProArgAlaLeuMetAlaP 260  
 DB 967 CGAATCGGTGACACGCTTGCGCTCACTCGCGCGCGCGCGCGCTCATGAGCGCC 1026

QY 260 cTyTrpTrpLysArgLeuGlyArgAspAlaLeuLeuSerTrpAspPheValLeuAlaValG 280  
 DB 1027 CTACTACAGAGGCTGGGCGCGCGCGCTCTCACTGAGGAGCAAGCGTGGCGGTGCA 1086

QY 280 nSerLeuTrpGlyLysProLeuGlyLysSerValAlaValGlnLeuProGlyLysLeuPhe 300  
 DB 1087 GAGCTGTATGGAGAACCCCTAGGGGCTCACTGTGCGCTCAAGCTCCAGAGAAAGCTTT 1146

QY 300 eThrAspPheGlnThrTrpAspSerTrpSerProGlnGlyArgArgProGlnTrpGlnG 320  
 DB 1147 CACTGACTTTGAGACTGTGAGACTCTTACAGCCCCCAAGAGAGCGCGCTGAAGCAGAGG 1206

QY 320 yProLysTrpCysHisSerSerPheAspAlaIleThrValAspArgGlnGlnLeuTrp 340  
 DB 1207 CCTAAATACGCACTCTCTTGTGATGCATCATGTATAGACAGGCAACAGCAACTGTATA 1266

QY 340 rIlePheLysGlySerHisPheTrpGlnValAlaAlaAspGlyAsnValSerGlnProAr 360  
 DB 1267 CATTTTAAAGGAGCATTTCTTGAGAGGTGGCAGCTGATGAGCAACGTCTCAGAGCCCG 1326

QY 360 gProLeuGlnGluArgTrpValGlyLeuProProAsnIleGlnAlaAlaValSerLe 380  
 DB 1327 TCCACTCAGAGAAAGATGGGTGCGGCTGCCCCCAACATGAGGCTGGGAGGTGACTT 1386

QY 380 uAsnAspGlyAspPheTrpPhePheLys----- 389  
 DB 1387 GAATGATGAGATTCTTACTTCTTCAAGTGCAATCCGTTTGAATTGATGATACCAAC 1446

QY 390 -----GlyGlyArgCysTrpArgPheArgGlyPr 399  
 DB 1447 CATCGTGGCTCCAGCTTAAAGAGACAGAGGGGGTGTGATGTGAGAGTTCCGCGGCC 1506

QY 399 cLysProValTrpGlyLeuProGlnLeuCysArgAlaGlyGlyLeuProArgHisProAs 419  
 DB 1507 CAAAGCATGTGGGTCTCCACAGCTGTGCGGCGAGGGGCTGCGCCGCAATCTGA 1566

QY 419 pAlaAlaLeuPhePheProProLeuArgArgLeuIleLeuPheLysGlyAlaArgTrpTy 439  
 DB 1567 GCGCGCTCTTCTTCTCTCTCTCTGCGCGCTCACTCTTCAAGGCTGCGCGCTACTA 1626

QY 439 rValLeuAlaArgGlyGlyLeuGlnValGluProTyTrpProArgSerLeuGlnAspTr 459  
 DB 1627 CGTCTGAGCCGAGAGGGGACTGCAAGTGAAGCTTACTACCCCGAAGTGCAGAGACTG 1686

QY 459 pGlyGlyIleProGlnGlnValSerGlyAlaLeuProArgProAspGlySerIleIlePh 479  
 DB 1687 GGGAGGATCCCTGAGAGAGGTGACAGCGCGCTGCGCAAGGCGCATGAGCTTCATCATCTT 1746

QY 479 ePheArgAspAspArgTrpTrpArgLeuAspGlnAlaLysLeuGlnAlaThrThrSerG 499  
 DB 1747 CTTCGAGATGACCGCTACTGTGCGCTCGACAGGCCAAATGAGCAACCACTCGCGG 1806

QY 499 yArgTrpAlaThrGlnLeuProTrpMetGlyCysTrpHisAlaAsnSerGlySerAlaLe 519  
 DB 1807 CCGCTGGGCGCAACGAGCTGCGCTGATGGGCTGTGCAATGCCAATCGGGAGCGCGCT 1866

QY 519 uPhe 520  
 DB 1867 GTTC 1870

RESULT 14  
 AR211784  
 LOCUS AR211784 2275 bp DNA linear PAT 20-JUN-2002  
 DEFINITION Sequence 1 from patent US 639371.  
 ACCESSION AR211784  
 VERSION AR211784.1 GI:21515198  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 2275)  
 AUTHORS Falduto,M.T., Magnuson,S.R. and Morgan,D.W.  
 TITLE Human matrix metalloproteinase gene, proteins encoded therefrom and  
 methods of using same  
 JOURNAL Patent: US 639371-A 1 04-JUN-2002;  
 FEATURES  
 source location/Qualifiers  
 1..2275  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,73e-110 Length: 2275  
 Score: 2764.00 Matches: 514  
 Percent Similarity: 99.23% Conservative: 3  
 Best Local Similarity: 98.66% Mismatches: 3  
 Query Match: 97.53% Indels: 2  
 Gaps: 0

|    |  |             |   |      |
|----|--|-------------|---|------|
| QY |  | 360         | gProLeuEnglUarXgtTTPValGlyLeuPProProAamIleGlValAlaAlaValSerle | 380  |
| Dd |  | 1125        | TCCACATGCAGAAAGATGGCTCGGACTGCCCCCCCAACATTGAAGCTCGCAGGTTCATT   |      |
| QY |  | 380         | uAnaBPglYaAPhePheTyYPhePheUySgiYgiYArCySTPaAgPhaRgiYProLy     | 1188 |
| Dd |  | 1185        | GAAATGATGAGATTTCTACTTCTTCAAAGGGGCTCATCTGAGAGTTCGGGGCCCCAA     |      |
| QY |  | 400         | sProValTTPGLyLeuPProGlnLeuCyBaARgAlaGlyLeuPProARghIsProaApAl  | 1244 |
| Dd |  | 1245        | GCCAGTGtGGGCTCTCCCAAGCTGTGCCGGCAAGGGGGCCCTGCCCGCCCATCTGAAGC   |      |
| QY |  | 420         | aAlaLeuPhePheProProLeuMaRgaRgeuLileuPheUySgiYAlaRTTYrVa       | 1304 |
| Dd |  | 1305        | CGCCCTCTTCTCCCTCCCTGTCGGCCGCTCATCTTTCAGAGGTGCCCGCATCACCT      |      |
| QY |  | 440         | lleuAlaARgYgiYgiYLeuGlnValIGluProTyTYProARghSerleuGlnAsPTPgl  | 1364 |
| Dd |  | 1365        | GCTGGCCCAAGAGGGAGACTGCAAGTGAAGCCCTTACCTCCCGAAGCTTCACAAGAATGGG | 1424 |
| QY |  | 460         | ygiYliePProGlnGlnValISerGIYAAlaUPProAgProaBPglYserIleIlePhePh | 1484 |
| Dd |  | 1425        | AAGCATTCCTTAGAGAGGTGACGGCGCCCTGCGCAAGCCCGATRGCTCATCATCTTCTT   |      |
| QY |  | 480         | eARgAsPaSPaRgyTYrTPaRgaLeuABpGlnAlaUySeuGlnAlaThTrhSerGIYAR   | 1544 |
| Dd |  | 1485        | CCGAGATACCGCATCTGCGGCTTCACACAGGCCAAAATGACAGCCAACCACTTCGGGCG   |      |
| QY |  | 500         | gTPAlaThhGluLeuPProTTPMeGIYCysTRPHIsAlaAnSerGIYserAlaleuPh    | 1604 |
| Dd |  | 1545        | CTGGGCCACCGAGCTGCTCGATGGCGGTGCTGGCATGCAACTCGGGGAGCGCTGT       |      |
| QY |  | 520 e 520   |   |      |
| Dd |  | 1605 c 1605 |   |      |

RESULT 15  
COB42403

|            |   |             |     |        |                 |
|------------|---|-------------|-----|--------|-----------------|
| LOCUS      | COB42403  | 2336 bp     | DNA | linear | PAT 02-AUG-2004 |
| DEFINITION | Sequence 1050 from Patent EPI440981.  |             |     |        |                 |
| ACCESSION  | COB42403  |             |     |        |                 |
| VERSION    | COB42403.1  | GI:50894190 |     |        |                 |
| KEYWORDS   |   |             |     |        |                 |
| SOURCE     | Homo sapiens (human)  |             |     |        |                 |
| ORGANISM   | Homo sapiens (human)  |             |     |        |                 |
| REFERENCE  | KumarVotla; Metazoa; Chordata; Craniata; Vertebrata; Euarchontom,   |             |     |        |                 |
| AUTHORS    | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |             |     |        |                 |
| TITLE      | Ilogai,T., Sugiyama,T., Oeuki,T., Wakamasu,A., Sato,H., Ishii,S., Yamamoto,J., Isono,Y., Nagai,K. and Irie,R. |             |     |        |                 |
| JOURNAL    | Folia length human cdna   |             |     |        |                 |
| FEATURES   | Patent: EP 1440981-A 1050 28-JUL-2004;  |             |     |        |                 |
| source     | Research Association for Biotechnology (JP)   |             |     |        |                 |
|            | location/Qualifiers   |             |     |        |                 |
|            | 1..2336   |             |     |        |                 |
|            | /organism="Homo sapiens"  |             |     |        |                 |
|            | /mol_type="unassigned DNA"  |             |     |        |                 |
|            | /db_xref="taxon:9606"   |             |     |        |                 |

Alignment Scores:

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 1,96e-108 | Length:       | 2336 |
| Score:                 | 2721.00   | Matches:      | 508  |
| Percent Similarity:    | 97.70%    | Conservative: | 1    |
| Best Local Similarity: | 97.50%    | Mismatch:     | 1    |
| Query Match:           | 96.01%    | Indels:       | 12   |
| DB:                    | 6         | Gaps:         | 1    |

US-10-791-980-6 (1-520) x COB42403 (1-2336)  
QY 1 MetValAlaArgValGlyT...



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - nucleic search, using frame\_plus\_p2n model

Run on: June 14, 2005, 07:21:52 ; Search time 4896 Seconds  
(without alignments)  
4042.774 Million cell updates/sec

Title: US-10-791-980-6  
Perfect score: 2834  
Sequence: 1 MVARVGLLRALQLLMGLT.....WATLPMWCMHANSGLALF 520

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO/epool/US10791980/runat\_13062005\_104719\_19930/app\_query.fasta\_1.711  
-DB=EST -QFMT=faaetp -SUFFIX=est -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10791980 @CGN 1.1 4352 @runat\_13062005\_104719\_19930 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: gb\_eest1:\*  
2: gb\_eest2:\*  
3: gb\_hc:\*  
4: gb\_eest3:\*  
5: gb\_eest4:\*  
6: gb\_eest5:\*  
7: gb\_eest6:\*  
8: gb\_g981:\*  
9: gb\_g982:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description          |
|------------|--------|-------------|--------|----|----------------------|
| 1          | 2519   | 88.9        | 1658   | 3  | CR590246 full-length |
| 2          | 2142   | 75.6        | 3440   | 3  | AK051258 Mus muscu   |
| 3          | 1420   | 50.1        | 822    | 4  | BI914743 603184116   |
| 4          | 1385.5 | 48.9        | 911    | 4  | BM009001 603618834   |
| 5          | 1328   | 46.9        | 962    | 5  | BK329088 BX329088    |
| 6          | 1169   | 41.2        | 1010   | 5  | BK349007 BX349007    |
| 7          | 1089   | 38.4        | 1086   | 4  | BMS44516 AGENCOURT   |
| 8          | 1027.5 | 36.3        | 619    | 7  | CV026222 4093 full   |
| 9          | 1010.5 | 35.7        | 587    | 5  | BO694961 1001104 H   |

|   |    |       |      |      |   |          |            |
|---|----|-------|------|------|---|----------|------------|
| c | 11 | 986.5 | 34.8 | 1065 | 4 | BMS60236 | AGENCOURT  |
|   | 12 | 899   | 31.7 | 638  | 2 | AW001264 | mus26h03.x |
|   | 13 | 869   | 30.7 | 499  | 5 | BK280995 | BK280995   |
|   | 14 | 846   | 29.9 | 958  | 5 | BK360790 | BK360790   |
|   | 15 | 833.5 | 29.4 | 581  | 5 | BP325923 | BP325923   |
|   | 16 | 821   | 29.0 | 1310 | 5 | BO722257 | AGENCOURT  |
|   | 17 | 810   | 28.6 | 507  | 4 | BI976357 | BI976357   |
|   | 18 | 769.5 | 27.7 | 503  | 1 | AV617699 | AV617699   |
| c | 19 | 726   | 27.2 | 825  | 1 | AL544270 | AL544270   |
|   | 20 | 720   | 25.4 | 563  | 2 | AW519125 | AW519125   |
|   | 21 | 714   | 25.2 | 822  | 5 | BU232388 | BU232388   |
|   | 22 | 695.5 | 24.5 | 515  | 2 | BE722401 | BE722401   |
|   | 23 | 654   | 23.1 | 803  | 5 | BU446082 | BU446082   |
|   | 24 | 652   | 23.0 | 856  | 4 | BG722422 | BG722422   |
|   | 25 | 602   | 21.2 | 562  | 7 | CV030096 | CV030096   |
|   | 26 | 587   | 20.7 | 4268 | 3 | BC036495 | BC036495   |
| c | 27 | 582.5 | 20.6 | 694  | 6 | BY743962 | BY743962   |
|   | 28 | 581.5 | 20.5 | 997  | 5 | BK360789 | BK360789   |
|   | 29 | 576   | 20.3 | 1045 | 2 | BBS31280 | BBS31280   |
|   | 30 | 570   | 20.1 | 4346 | 3 | BC047278 | BC047278   |
|   | 31 | 568.5 | 20.1 | 883  | 4 | BT756398 | BT756398   |
|   | 32 | 559.5 | 19.7 | 662  | 7 | CO506713 | CO506713   |
|   | 33 | 555   | 19.6 | 1749 | 9 | AT421188 | AT421188   |
|   | 34 | 553.5 | 19.5 | 850  | 4 | BI559993 | BI559993   |
|   | 35 | 549.5 | 19.4 | 2238 | 3 | AK079924 | AK079924   |
|   | 36 | 547   | 19.3 | 601  | 2 | BB660215 | BB660215   |
|   | 37 | 540.5 | 19.1 | 1749 | 9 | AY421190 | AY421190   |
|   | 38 | 540.5 | 19.1 | 2543 | 3 | AK088476 | AK088476   |
|   | 39 | 536   | 18.9 | 1135 | 2 | BE736871 | BE736871   |
| c | 40 | 534.5 | 18.9 | 465  | 1 | AI816763 | AI816763   |
|   | 41 | 530   | 18.7 | 375  | 6 | CB691626 | CB691626   |
|   | 42 | 521.5 | 18.4 | 2204 | 3 | AK086808 | AK086808   |
|   | 43 | 521.5 | 18.4 | 2438 | 3 | BC045610 | BC045610   |
|   | 44 | 521.5 | 18.4 | 2438 | 3 | BC045610 | BC045610   |
|   | 45 | 516   | 18.2 | 2267 | 3 | BC078136 | BC078136   |

## ALIGNMENTS

RESULT 1  
CR590246 1658 bp mRNA linear HTC 21-JUL-2004  
LOCUS full-length cDNA clone CSDID075Y06 of Placenta Cot 25-normalized  
DEFINITION of Homo sapiens (human).  
ACCESSION CR590246.1 GI:50471053  
VERSION CR590246.1  
KEYWORDS HTC; CDS; cDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
Paradey Avenue  
2 (bases 1 to 1658)  
REFERENCE Genoscope.  
AUTHORS Direct Submission  
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail: sequef@genoscope.cns.fr

COMMENT - Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
FEATURES Location/Qualifiers  
1..1658  
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSD01075YGC6"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"
Cot 25-normalized

```

## US-10-791-980-6 (1-520) x CR590246 (1-1666)

|           |               |      |
|-----------|---------------|------|
| 1.94e-185 | Length:       | 1658 |
| 2519.00   | Matches:      | 469  |
| 99.79%    | Conservative: | 0    |
| 99.79%    | Mismatches:   | 1    |
| 88.88%    | Indels:       | 1    |
| 3         | Gaps:         | 0    |

1-1658) CR390246

1038 ————— **Fr**

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE PUBMED 20530913 11076861

REFERENCE AUTHORS 4

TITLE The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

JOURNAL NATURE 409, 685-690 (2001)

REFERENCE AUTHORS 5

TITLE The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

JOURNAL NATURE 420, 563-573 (2002)

REFERENCE AUTHORS 6 (bases 1 to 3440)

ADACHI, J., ALZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FURUKI, M., HANAGAKI, T., HARA, A., HASHIZUME, W., HAYASHIDA, K., HAYATEU, N., HIRAMOTO, K., HIRAKA, T., HIROZANE, T., HOJI, F., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KAKUKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KOUNDA, M., KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M., NAKAMURA, Y., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHSATO, N., OKAZAKI, Y., SAIJO, R., SAIJOH, H., SAKAI, C., SAKAI, K., SAKAZUME, N., SANO, H., SASEKI, D., SHIBATA, K., SHINGAWA, A., SHIRAKI, T., SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKIHARA, S., TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A., MURAMATSU, M. and HAYASHIZAKI, Y.

Direct Submission

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Shuohito-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://lantom.gsc.riken.jp/

Location/Qualifiers

1..3440

organism="Mus musculus"

molecule="mRNA"

strain="CS7BL/6J"

db\_xref="PANTOM.DB:D130023P12"

db\_xref="taxon:10090"

clone="D130023P12"

tissue\_type="spinal ganglion"

clone\_lib="RIKEN full-length enriched mouse cDNA library"

dev\_stage="12 days embryo"

167..1657

/note="unnamed protein product; matrix metalloproteinase 28 (epilysin) (MGP)|MGI:2153062, GB|AF36345, evidence: BLASTN, 99%, match=982)

putative"

codon\_start=1

protein\_id="BAC34577.1"

db\_xref="GI:26341830"

/translation="WVAGVSLILRALPLLWGCDQAPTOHGLPELRQEAARFLEKYG YLSEQSAAPASQAFRNAIREFQWISQPLSGVLDQATLRQWTRRCVAATDSHATV TERISTLAGRAKKRRKRPFAKPNKNGKHLRYLVNWPBEPVRAFAFO LMSVSALEPWPAPATGPADIRLPFGDDHNDLANADPGGLAAHAFVLRGGAH DGDERSLSRRGRNLFVYLAHEIGTIGLTHSPAPRLMAPYTKLDRDLISDDV LAVQSLYKPLGRVATQPLGVFTDFEAMPDHSQSRPETRGKYSDFDATIVE PRLQKRWPGILPGIEAAVSLIEDGPFYFKGNRCWRPGTKTSVWGFAQLCRAGLPR

HPDAALEPPPIRLVLFEKGRYYVLAQGMQVEYYPRSLDMAGVEEVSAGLPRPD GSIIFFRDHWHMDQAKLRTVSSGRATETLSMGMNANGALF"

ORIGIN

Alignment Scores:

Pred. No.: 1,156-155 Length: 3440

Score: 2142.00 Matches: 402

Percent Similarity: 82.73% Conservative: 29

Best Local Similarity: 77.16% Mismatches: 65

Query Match: 75.58% Indels: 26

DB: 3 Gaps: 1

US-10-791-980-6 (1-520) x AK051258 (1-3440)

QY 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuTrpGlyYHisLeu 20

DB 167 ATGGTGGCTGGCGGTGACGTCTCTGTGGCCGCCCTGCGCGCTGTGCTGGGGCGCCAG 226

QY 21 AspAlaGlnProAlaGlnArgGlyGlyGlnGlnLeuArgValGlnAlaGlnAlaPheLeu 40

DB 227 GACGACAGCCACCCACAGCAGACTCCCGAGCTGCCAGAGGACAGGACATTCCTG 286

QY 41 GlnuYrTrpGlyYrTrpLeuAsnGlnGlnAlaProlysaAlaProThrSerThrArgPheSer 60

DB 287 GAGAAgTrpGATACCTCAGTGAAGAGGAGGCTCCAAAGCCAGCCTCCGACAGTTCAAG 346

QY 61 AspAlaIleArgAlaPheGlnTrpValSerGlnLeuProValSerGlyValLeuAspArg 80

DB 347 AACGCCATCAGAGATTCATGATGATATCCAGCTGCCCTTCAGTGGTGTGTCGACAG 406

QY 81 AlaThrLeuArgGlnMetThrArgProArgCysGlyValTrpAspThrAsnSerTyraAla 100

DB 407 GCCACACTGCCCGAGATGACCCGCCAGCTGGGGGTGGAGATCAGACAGTCATGCC 466

QY 101 AlaTrpAlaGlnuArgGlnLeuSerAspLeuPheAlaArgHisaArgThrIlyMetArgArgIys 120

DB 467 ACTTGACAGAGAGATACGATACCTGTCTTCGACCCGAGCCAAATAGAGGCGTAAAG 526

QY 121 LysArgPheAlaIysGlnGlyAsnIlySerTyraIysGlnHisIleuSerTyraArgLeuVal 140

DB 527 AACGCTTGCAAAACCGATGACGATGACAAATGATGACAAAGCATCTTCTTACCGCTGGTG 586

QY 141 AsnTrpProGlnHisIleuArgSerArgGlnPheGlyAlaProCysAlaProProSerSer 160

DB 587 AACTGGCTGAGCGGCTCCGAGCAGCAGCTGTGCAAGGGGCGGTGCGCTTCAG 646

QY 160 rCysGlyAlaThrSerGlnArgTrpSerSerGlyArgProGlnProGlnAlaProLeuTh 180

DB 647 CTGGAGACTAAATGTGTCAAGCACTGAGATTCGGGAAGCCCAAGCCACGCGCTGCCAGC 706

QY 180 rSerGlySerProSerSerIysGlyThrThrThrMetGlyTrpAlaMetProLeuMetAl 200

DB 707 ATCCGCTTACCTTCTTCCAGAGACACCAATGATGAGGCTGGCCACGCTTCGATGGC 766

QY 200 aglnGlyAlaProTrpArgTrpProPheLeuProArgArgIysGlnuAlaHisIleAspArgI 220

DB 767 CGAGGGGGTGGCCGTGACACAGCCTTCTG -CCCCGGGGGGGCAAGGCATTTCAGCGG 825

QY 220 nAspGlnuArgTrpSerLeuSerArgArgArgGlyArgAsnLeuPheValIleuAlaHis 240

DB 826 AGACAGAGGCGTGTCTTGAACCCCGCGGGGGCAACTGTTCGTGGTGGTGGCCCA 885

QY 240 gGlnIleGlyHisIleThrLeuGlnGlyLeuThrHisSerProAlaProArgAlaLeuMetAlPr 260

DB 886 CGAATTCGCGCACAGCTCGGTCTCAGCATCTCGCCCGCGCACGGGGCTTATGCGCC 945

QY 260 oTyTrpIlysaArgLeuGlyArgAspAlaLeuLeuSerTrpAspArgValLeuAlaValG 280

DB 946 CTACAGCAAGAAGCTGGGCGCGACGCTTCTCAGCTGAGGACGATGTGCGCGGTGCA 1005

QY 280 nSerLeuTrpGlyYrTrpProLeuGlnGlyYrSerValAlaValGlnLeuProGlyYrLeuPh 300

DB 1006 GAGCTGTACGGAAGCCTTCGGGACGCTCAGTAGCCACCAACTCCCCGGAAGGTGT 1065

```

QY      300  ethrAaphegIuThrTpAaseryrSerProGInGlyArGArProGluThrGInG 320
Db      1066  CACTGACTTGAAGGCTGGAGACCCCAACTCCAGACAGACAGCCCGGAAACAGAGG 1125
QY      320  yProLyTyCyehShSerSerPheAapAlaIeTrValaAapArGInGInGLeuTy 340
Db      1126  TCCTAAATACGACCTCTTCTTGAATGACATCACTGA----- 1165
QY      340  rIlehelYeglySerHisPheTrpGluValAlaAlaAapGlyAenValSerGluProAr 360
Db      1166  -----GAGCTCG 1173
QY      360  gProLeuGInGlyArGTrpValGlyLeuProProAanIleGluAlaAlaValaSerle 380
Db      1174  TCACATACAGAAAGATGGCCAGAGACTGCCACTGGATTCAGGCTCTCAAGTCTCTT 1233
QY      380  uAanAapGlyAapPheTyPhePheLyGlyGlyArGyETrPArGArGlyProly 400
Db      1234  GGAAGATGAGACTTCTACTTCTTAAAGGAATCGTGTGGAGGTTCCAGGCGACGAA 1293
QY      400  sProValTrpGlyLeuProGInLeuCyArGAlaGlyGlyLeuProArGHisProAapAl 420
Db      1294  GTCCGCTGTGGGGTTCCGACAGCTTCTCCGGGAGGTTGCTAACCCGACACCCGAGCG 1353
QY      420  aAlaLeuPhePheProProLeuArGArGlyLeuPheLyGlyAlaAapGlyTyTyVa 440
Db      1354  AGCTCTCTTCTTCCCGCTCTGCGCGCCCTAGTCTTCAAGGGTTCCGCTACTAGT 1413
QY      440  lLeuAlaArGlyGlyLeuGInValGInProTyTyTyTyProArSerLeuGInAapTrpG 460
Db      1414  GCTAGCTCAAGAGAGGATGCAAGTGAAGCCCTACTATCCCGCAGCCTGGCAGACTGGG 1473
QY      460  yGlylLePProGluGluValSerGlyAlaLeuProArGProAapGlySerlelelePhe 480
Db      1474  CGGTGTCCCGAGAGAGTCAAGCGCGCTGCCAAGGCGGATGCTCCATATCTCTT 1533
QY      480  eArGAPAPAPArGTrPArGlyLeuAapGInAlaLyLeuGInAlaThrThSerGlyAr 500
Db      1534  CAGAGATGACCACTATTGGCATTTGGATCAAGCCAAAGCTGGGTGACACCTCCGGT 1593
QY      500  gTrPAlaThrGlyLeuProTrpMetGlyCyETrPHIsAlaAasSerGlySerAlaLeuPh 520
Db      1594  CTGGGCTACCGAGGCTGCTGGATGGGCTGTGGAAATGCCAACTCAGAGAGTCCCTGT 1653
QY      520  e 520
Db      1654  C 1654

```

http://image.llnl.gov  
 Plate: LLM1626 row: 0 column: 04  
 High quality sequence stop: 821.  
 Location/Qualifiers  
 1. 822  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5248347"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPOrt6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."

Alignment Scores:  
 Pred. No.: 2,54e-100 Length: 822  
 Score: 1420.00 Matches: 268  
 Percent Similarity: 97.45% Conservative: 0  
 Best Local Similarity: 97.45% Mismatches: 5  
 Query Match: 50.11% Indels: 5  
 DB: 4 Gaps: 0

US-10-791-980-6 (1-520) x B1914743 (1-822)

```

QY      119  ArgLyLyArPheAlaLyGInGlyAaNLyETrTyTyTyTyTyTyTyTyTyTyTyTyTy 138
Db      2  CCTAAGAAACGCTTTGCAAGCAAGTAAACAAATGTACAGACGACCTTCTCAAGCG 61
QY      139  LeuValaEntPProGluThrSleu-ArSerArGInPheGlyAlaProCyAlaProPr 158
Db      62  CTGGTGAATCGGCTTGATGATCTGCGGAGCGGAGCTTGGGGGCGGTCGCGCGCC 121
QY      158  oSerSerCyGlyAlaThrSerGInArGTrPserSerGlyArGProGInProGInAlaPr 178
Db      122  TTCAGATTGTGGAGCAACCTTCAGGCTGAGGTTCTGGAGGCCGCCAGCCAGCGCC 181
QY      178  OLeuThrSerGlySerProSerSerLyGlyThrThrThrMetGlyTrpAlaMetProle 198
Db      182  GCTGACATCCGGCTCACTTCTTCCAGGGAGCAACCAATGGAGGCTGGCAATGCTCT 241
QY      198  uMetAlaGInGlyAlaProTrPArGTrPProPheLeuProArGArGlyGlyAlaHisPh 218
Db      242  GATGGCCCAAGGGGGGCGCTGGCCACGCTTC-CTGCCCGCGCGGGAAGCGGACTT 300
QY      218  eAapGInAapGluArGTrPserLeuSerArGArGlyArGAsnLeuPheValle 238
Db      301  CGACCAAGATGAGCGCTGGTCCCTGAGCCGCCCGCGGCGCAACCTGTGTGGTCT 360
QY      238  uAlaHisGluIleGlyHisThrLeuGlyLeuThrHisSerProAlaProArGAlaLeu 258
Db      361  GCGCGAGAGATGGGTCAACAGCTTGCTTACCCATCACCAGCGCGCGCGGCTCAT 420
QY      258  tAlaProTyTyTyLyArGLeuGlyArGAPAlaLeuLeuSerTrpAapPValleuAl 278
Db      421  GCGCCCTTACTTCAAGAGCTGGGCGCGAGCGCTGTCTGAGTGGGACAGCTGCTGGC 480
QY      278  vAlaGInSerLeuTyTyTyLyPProLeuGlyGlySerValAlaAlaGInleuProGly 298
Db      481  CGTGCAAGCTGTATGGGAAGCCCTTGGGGGCTCACTGAGCGGTCCAGCTCCAGGAA 540
QY      298  sLeuPheThrAapPheGluThrTrpAapSerTySerProGInGlyArGArProGluTh 318
Db      541  GCTGTTCAGTCTTGAAGCTGGAGCTCTTACAGCCGCCAAGAGAGCGCTGAAAC 600

```

RESULT 3  
 B1914743  
 LOCUS 822 bp mRNA linear EST 17-OCT-2001  
 DEFINITION 60384116F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5248347 5',  
 mRNA sequence.  
 B1914743  
 VERSION B1914743.1 GI:16199027  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 Unpublished (1999)  
 Contact: Robert Strauberg, Ph.D.  
 Email: rgs@bbs-riemall.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT



Qy 338 tGInglYProLYtYrCySHIsSerSeRheAsPaLaIeThrVaLaSPaRGInGlInGl 338  
 Db 601 GCAGGCCCCCTTAATACTGCCACTTCTTCGATGCCATCAGTGAACAGGCAACAGCA 660  
 Qy 338 nLeuYrTLePheLYSeRHeThIePhETrPGIuValAlaLaSPaRGLYaSPVaISerGI 358  
 Db 661 ACTGTAACATTTTACAGGAGCCATTTCTGGAGGCTGAGCTGATGCGAACGCTTCAGAA 720  
 Qy 358 u-PrArGPProLeuGInGluaRGTrPVaIGLYLeuPProPaSnnIeGIuAlaLaIaY 378  
 Db 721 ACCCGCTTCATCTGACGAAAGATGGCTGGGCTGCCCAACATTGAGGCTGCGGCA 780  
 Qy 378 aISerLeuAaPaRGLYaSPHeTYrPhePheLYSeGYLY 391  
 Db 781 TGTCAATTGACTGATGAGATTCTTACTTCTTCAAGGGGT 821  
 Db  
 RESULT 4  
 BM009001 911 bp mRNA linear EST 30-OCT-2001  
 LOCUS 603618834P1 NIH\_MGC\_39 Homo sapiens cDNA clone IMAGE:5422808 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BM009001  
 VERSION BM009001.1 GI:16523355  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 911)  
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgarber@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
 http://image.llnl.gov  
 Plate: L1CM1881 row: d column: 09  
 High quality sequence scop: 767.  
 FEATURES  
 source  
 1.911  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5422808"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_11b="NIH\_MGC\_39"  
 /note="Organ: pancreas; Vector: pOTB7, Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed  
 by Ling Hong in the laboratory of Gerald M. Rubin  
 (University of California, Berkeley) using ZAP-cDNA  
 synthesis kit (Stratagene) and Superscript II RT (Life  
 Technologies)."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.19e-97 Length: 911  
 Score: 1386.50 Matches: 281  
 Percent Similarity: 92.53% Conservative: 4  
 Best Local Similarity: 91.23% Mismatches: 12  
 Query Match: 48.92% Indels: 12  
 DB: 4 Gaps: 3  
 US-10-791-980-6 (1-520) x BM009001 (1-911)  
 Qy 109 LeuPheAlaARHIsaRThrLYs-MeTaRGaRGLYeLYSaRPhaLaLYaSGInGlY 128

Db 3 TGTGTTGCTAGACACCGGACCAAGATAGGCGGTGAAGAAAGCTTTGGCAAGAGTA 62  
 Qy 128 nLYeATrPYrLYSeGInHIsLeuSeTYrYAGYLeuValAsnTrPProGInHIsLeu-Arg 147  
 Db 63 ACAAAATGGTACAGAGGACCTCTCTACCGCTGTGAATCGGCTTGAAGCTTGGCGG 122  
 Qy 148 SeTaRGInPheGlyAlaProCySaLaProPSeSerSeCYSeGYAlaThSeRGInaRG 167  
 Db 123 AGCGGAGATTGGGGCGCGGTGGCGGCGCTTCAAGTTGGAGCAACGCTCAGGCG 182  
 Qy 168 TrPSeSerSeGYLYaRGProGInPProGInAlaProLeuThrSeRGLYSeTrPSeSerLYs 187  
 Db 183 TGAAGTTCTGGAGGAGCGCCAGCACAGGCGCGCTGAATCGGCTCACTTCTTCAG 242  
 Qy 188 GLYThTrThrThMeGlyTTPAlaMeRProLeuMeTaIaGInGlyAlaProTrPaRGTr 207  
 Db 243 GGGACCAACACATGAGGCTGGGCAATGCTTGAATGGCCAGGAGGCGGCTGTGCAGC 302  
 Qy 208 ProPheLeuProARGaRGLYGlyuAlaHIsPheARGLnaSPGLuaRGTrPSeSerLeuSer 227  
 Db 303 CCTTC-CTGCGCGCGCGCGGCGGAGCGCACTTCAGCAAGATGAGCGGCTGCTGAGC 361  
 Qy 228 ARGaRGaRGLYaRGaSPaSPaPheValaLeuAlaHIsGlyuIleGlyHIsThrLeuGly 247  
 Db 362 CGCGCGCGCGCGCAACCTGTTCTGGTGGCGGCGAGATCGGTCTACACGCTTGCC 421  
 Qy 248 LeuThrHIsaSPProAlaProARGLaLeuMeTaLaProTYrTYrLYaARGLuGlyLYaRG 267  
 Db 422 CTCACCACTGCGCGCGCGCGCGGCTGATGAGGCTTATACAGAGGCTGGAGGCC 481  
 Qy 268 AsPaLaLeuLeuSeTrTPaSPaSPaValaLeuAlaGInSeRLeuTYrGLYs-ProLe 287  
 Db 482 GACGGGCTGCTAGCTGGGAGCGAGCTGGCGCGTGAAGCTGTATGGAGGCCCT 541  
 Qy 287 uGlyLYSeRValaLaValGInLeuPProGlyLYLeuPheThRaSPheGlyThTPa 307  
 Db 542 AGGGGCTCAGTGGCGCTGCCAGCTCCAGAAAGCTGTCACTTGAGACCTGGGA 601  
 Qy 307 pSeTYrSeRProGInGlyARGaRGProGInThrGInGlyProLYtYrCySHIsSeRSe 327  
 Db 602 CTCCTACAGCCCCCAAGAGGCGCTGAAACGCAAGGCGCTTAATCTGCACCTTC 661  
 Qy 327 rPheAsPaLaIeThrVaLaSPaRGInGInLeuTYrTLePheLYSeRHeThIeP 347  
 Db 662 CTTCATGCTCATCAGTGAACAGGCAACAGAACTTGTACTTTTACCGGAGCACT 721  
 Qy 347 heTrPGIuValAlaLaSPaRGLYaSPVaISerGIu-ProARGPProLeuGInGlu-ArgTr 366  
 Db 722 TCTGGAGAGTGGCACTGATGCAAGCTCAGAGCCCGTCCACTGACAGGAAAGACG 781  
 Qy 366 pValGlyLeuPProPaSnnIeGIuAlaAlaLaValaSeRLeuAaSPaRG 383  
 Db 782 GGTGGAGACTGCGCCCAAAATGAGAGCTTGGGCGAGATGATATTCAGTCACTGG 841  
 Qy 383 YaSPHeTYrPhe-PheLYSeGYLYaRGYsTrPaRGPhaRGLYProLYsProValT 403  
 Db 842 CGAATTCATCTTCTTCAAGGGGGGTCAATGTGAAGTTCCGGGGCCCAAGCACTGT 901  
 Qy 403 TrPGLYLeu 405  
 Db 902 GGGGTCTC 909  
 RESULT 5  
 BX329088/c 962 bp mRNA linear EST 07-APR-2004  
 LOCUS BX329088 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS001075YG06 3-PRIME, mRNA sequence.  
 ACCESSION BX329088  
 VERSION BX329088.2 GI:46269851  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
AUTHORS  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished (2001)  
COMMENT  
On May 2, 2003 this sequence version replaced gi:30334822.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by life technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8754.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0BA1028ZC01\_CS02595\_1&c=8754.r

FEATURES  
source  
Location/Qualifiers  
1..962  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1075YG06"  
/cissue="CS0D1075YG06"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized." sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

## Alignment Scores:

Pred. No.: 4 63e-93  
Score: 1328.00 Length: 962  
Percent Similarity: 88.04% Matches: 275  
Best Local Similarity: 84.36% Conservative: 12  
Query Match: 46.86% Mismatches: 21  
DB: 5 Indels: 19  
Gaps: 6

US-10-791-980-6 (1-520) x BX329088 (1-962)

QY 160 SerCyGgIyAlaThSerGlnAgtTyrSerSer-----GlyArgProGln--- 174  
DB 960 AGTTGTGG-----AACCAGAGTTTCCAGCGGTGGAGTTTGGAGAGCCCAAGCC 907  
QY 175 ProGlnAlaProLeuThr---SerGlySerProSer---SerLyGlyThrThrThMet 192  
DB 906 CCAAGGCCCCCGCTGAACATTCGGGCTCACCCTTCCCAAGGAGGACCAACAGCAAT 847  
QY 193 GlyTTrpAla-MetProLeuMet-AlaGln-GlyAlaProTTrpAgtThrProPheLeuPro 211  
DB 846 GGGCTGGGCAAAATGCTTTATGAGGCGCCAGGCGGCGCTTGGCCAGCCCTTCATGCCCC 787  
QY 212 ArgArg-GlyGlu-AlaHisPheAer-GlnAerGlu-AgtTTrSerLeuSer---ArgAr 229  
DB 786 CCGCCCGGCGAAGCGGCACTTCGAACCAAAATAGAGCGTGGTCCCGAAGCCCGCCG 727  
QY 229 gArGgIyArGanLeuPheValAlaAlaHisGluIleGlyHisThPheLeuTh 249  
DB 726 CCGCGGGGCAACCGTTTCGTGCTGCGCGCAGAGATCGGTCACAGATTGGCTCAC 667  
QY 249 r-HisSerProAlaProArgAlaLeuMetAlaProTTrpTyrTyrArgGluGlyArgAspA 269  
DB 666 CCAACTCGCCCGCGCGCGCGCTCATGCGCCCTACACAGAGGCTGGCGCGCAGC 607  
QY 269 lAlaLeuSerTTrpAspAspValLeuAlaValGlnSerLeuTyrGlyLyProLeuGlyG 289  
DB 606 CGCGTCTCACTGAGCAGACGTCTGCGCGGTGCAAGCCTGTATGGAGNCCCTTAGGGG 547  
QY 289 lYserValAlaValGlnLeuProGlyLyLeuPheThrAspPheGluThrTTrpAspSerT 309

DB 546 NCTCAGTGGCCGTCACAGCTCCAGAAAGCTGTTCACCTTGGAGACTGGACTCTCT 487  
QY 309 YrSerProGlnGlyArgArgProGluThrGlnGlyProLySerTyrCyHisSerSerPhea 329  
DB 486 ACAGCCCCCAAGAGAGAGCGCCCTGAAACGACAGGCGCTTAATATCTGCCACTTCTTCG 427  
QY 329 sPaAlaIleThrValAaPargInGlnGlnLeuTyrIlePheLyGlySerHisPheTrpG 349  
DB 426 ATGCATCACTGTAGACAGGCAACAGCAACTGTACATTTTAAAGGAGCATTTCTGGG 367  
QY 349 luvAlAlaAlaAerGlyAenValSerGluProAgtProLeuGlnGlyAgtTrpValGly 369  
DB 366 AGGTGCACTGATGAGCAACGCTTCAGAGCCCGCTCCACGTGAGAAAGATGGGTGGGC 307  
QY 369 euProPheAmnIleGluAlaAlaValSerLeuAenAerGlyAerPheTyrPhePheL 389  
DB 306 TGCCCCCAACATTAAGGCTGCGGACAGTGCATTGATATAGATTTCTACTTCTTCA 247  
QY 389 yGgIyGlyArgCySerTrpArgPheArgGlyProLySerProValTTrpGlyLeuProGlnLeu 409  
DB 246 AAGGGGGTGCATGCTGAGAGGTACCGGGAAACAGCACTGTGGGGTCTCCACAGCTGT 187  
QY 409 YrArgAlaGlyGlyLeuProArgHisProAspAlaAlaLeuPhePheProLeuArgA 429  
DB 186 GCGCGGAGGGGGCTGCCCCCGCATCTGCGCGCCGCTTCTTCCCTCTCTTGGGC 127  
QY 429 rGluIleLeuPheLyGlyAlaArgTyrTyrValIleAlaArgGlyGlyLeuGlnValG 449  
DB 126 GCTCATCTCTTCAAGGCTGCGGCTACTACGAGTGGCGCCAGGGGAGACTGCAAGTG 67  
QY 449 luvProTTrpProArgSerLeuGlnAerTTrpGlyIleProGlnGluValSerGlyA 469  
DB 66 AGCCATCACTACCCCGAAGTCTGCAAGGAGTGGAGGAGCATCTCTTNAAGTTCGGCGG 8  
QY 469 la 469  
DB 7 CC 6

## FEATURES

source  
Location/Qualifiers  
1..1010  
/organism="Homo sapiens"  
/mol\_type="mRNA"

REFERENCE  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
AUTHORS  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished (2001)  
COMMENT  
On May 5, 2003 this sequence version replaced gi:30369337.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized. Library  
was constructed by life technologies, a division of Invitrogen. This  
sequence belongs to sequence cluster 8754.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0BA0372B01\_CS03458\_1&c=8754.r

Location/Qualifiers  
1..1010  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
 /clone="CS0D1075YG06"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

## Alignment Scores:

| Pred. No.:             | 1.19e-80 | Length:       | 1010 |
|------------------------|----------|---------------|------|
| Score:                 | 1169.00  | Matches:      | 264  |
| Percent Similarity:    | 81.12%   | Conservative: | 11   |
| Best Local Similarity: | 77.88%   | Mismatches:   | 49   |
| Query Match:           | 41.25%   | Indels:       | 18   |
| DB:                    | 5        | Gaps:         | 3    |

US-10-791-980-6 (1-520) x BX349007 (1-1010)

QY 136 SerTyrArgLeuValAsnTrpProGluHisLeu-ArgSerArgGlnPheGlyAlaProCys 155  
 Db 3 TCTTACCGCTGTGTAAGTGGCTGAGCATCTGCCGAGCCGCGAGTTCGGGGCCGCTTG 62  
 QY 155 GAlaProProSerSerCysGlyAlaThrSerGlnArgTrpSerSerGlyArgProGlnPr 175  
 Db 63 CGCGCCGCTTCAAGTTGTGGAGCAACGCTCAGCGCTGGAATTCGGAGGCCCAAGCC 122  
 QY 175 GGlAlaProLeuThrSerGlySerProSerSerGlyThrThrThrMerGlyTrpAl 195  
 Db 123 ACAGGCCCCGCGAGCATCGGCTCACTCTTCCAAAGGAGCAACAACATGGGCTGGGC 182  
 QY 195 AMeProLeuMetAlaGlnGlyAlaProTrpArgThrProPheLeuProArgArgGly 215  
 Db 183 AATGCTTTGATGGCCCCAGGGGGCCGCTGGCGCACGCTTC-CTGCCCGCGCGGCGA 241  
 QY 215 uAlaHisPheAspGlnAspGluArgTrpSerLeuSerArgArgArgGlyArgAsnLeuPh 235  
 Db 242 AGCGCACTTCCAGCAAGATGAGCGCTGGCTCTGAGCGCGCGCGCGGCGCAACTGTT 301  
 QY 235 eValValLeuAlaHisGluHisGlyHisThrLeuGlyLeuThrHisSerProAlaProAr 255  
 Db 302 CGTGCTGCTGGCGCAGAGATGGGTCAACGCTTGCCCTCACCACTGCCCGCGCGCG 361  
 QY 255 GAlaLeuMetAlaProTyrTrpLysArgLeuGlyValArgAspAlaLeuLeuSerTrpAspAs 275  
 Db 362 CGCGCTCATGGCGCTTACTACAAAGGCTGGGCGCGAGCGCTGCTCAGCTGGAGCA 421  
 QY 275 PValLeuAlaValGlnSerLeuTyrGlyLysProLeuGlyGlySerValAlaValGlnLe 295  
 Db 422 CGTGCTGGCGCGAGAGCTGTATGGAGAGCCCTAGGGGGCTCAGTGGCGCTCAAGC 481  
 QY 295 uProGlyLysLeuPheThrAspPheGluThrTrpAspSerTyrSerProGlnGlyArgAr 315  
 Db 482 TCCAGAAAGCTGTTCACACTTGAAGCTGGAGCTCTTACAGCCCCCAAGAGAGGGG 541  
 QY 315 GProGluThrGln-GlyProLysTyrCys-HisSerSer-PheAspAlaLeuThrValAs 334  
 Db 542 CCTTAAGAGCGAGGGGCGCTTAATAGTCCCACTCTTCCCTTGAAGGCCCATAGTGA 601  
 QY 334 P-ArgGlnGlnGln-LeuTyrIlePhe-LysGlySerHisPheTrp-GluValAlaAla- 352  
 Db 602 ACAGCAACAGCAACCTGTCATTTTAAAGGAGGACATTTCTGGGGAGGGGAGGCG 661  
 QY 353 AspGlyAsn---ValSerGluProArgProLeuGlnGluArgTrpVal-GlyLeuProPr 371  
 Db 662 GATGCAACCTTCTTAAGAGACCCCGTCCCTGCAAGAAATGGGTGCGTCCGCCCC 721  
 QY 371 GAsnIleGluAla---AlaAlaValSerLeuAsnAspGlyAspPheTyrPhePhe---Ly 389  
 Db 722 CAAATTTAGGCTTGGGCGAGGCGTCAATTGAAGAGAGGAATCTCTACCTTTTCAAAA 781  
 QY 389 gGlyGlyArgCys-TrpArgPheArgGlyProLysProValTrpGlyLeu-ProGlnLeu 408

Db 782 GAGGGGTGATGCTTGGAGAGTCCGGGGCCCCCANACCAAGGGGGGGTCTCCACAAATTG 841  
 QY 409 CysArgAlaGlyGlyLeuProArgHisProAspAlaAlaLeuPhePheProProLeuArg 428  
 Db 842 GTCCGGGAGGGGGGCGCTCCCGCCCATCTTA-CCGCGCTTTTCTTCTTTTGGC 900  
 QY 429 ArgLeuIleLeuPheLysGlyAlaArgTyrTrpValLeuAlaArgGlyGlyLeuGlnVal 448  
 Db 901 CGCCTAT--ATCTTTAAGGGTCCCGTTCATTAAATGCTGCGCCCGGGGAGCTGAATAG 958  
 QY 449 GluProTyrTrpProArgSerLeuGlnAspTrp 459  
 Db 959 GGCCCCCTTATCCCAATCTTGAAGGAGCGTGG 991

## RESULT 7

BMS44516 1086 bp mRNA linear EST 20-FEB-2002  
 LOCUS AGENCOURT\_6491854 NIH\_MGC\_124 Homo sapiens cDNA clone IMAGE:5728282  
 DEFINITION 5', mRNA sequence.

ACCESSION BMS44516  
 VERSION BMS44516.1 GI:18775843  
 KEYWORDS EST.

SOURCE  
 ORGANISM Homo sapiens (human)

REFERENCE  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bcr-mail.nih.gov  
 Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 .Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILMIL at:  
 http://image.llnl.gov  
 Place: LLM12723 row: 1 column: 11  
 High quality sequence start: 29  
 High quality sequence stop: 685.  
 Location/Qualifiers

## FEATURES

## source

1..1086  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5728282"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_124"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: EcoRV  
 (destroyed); Site: 2: NotI; RNA source male hippocampus,  
 age 27. Library is oligo-dT primed and directionally  
 cloned (EcoRV site is destroyed upon cloning). Average  
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 012."

## ORIGIN

## Alignment Scores:

| Pred. No.:             | 2.24e-74 | Length:       | 1086 |
|------------------------|----------|---------------|------|
| Score:                 | 1089.00  | Matches:      | 231  |
| Percent Similarity:    | 84.81%   | Conservative: | 9    |
| Best Local Similarity: | 81.63%   | Mismatches:   | 23   |
| Query Match:           | 38.43%   | Indels:       | 20   |
| DB:                    | 4        | Gaps:         | 6    |

US-10-791-980-6 (1-520) x BMS44516 (1-1086)

QY 1 MetValAlaArgValGlyLeuLeuLeuArgAlaLeuGlnLeuLeuTrpGlyHisLeu 20



Db 361 AACCCCTTGCAACCAAGTAACTAGTACAGCAGACCTCTCTACCGCTGGG 420  
 Qy 141 AentrpProgluHleu-ArgerSerArglnPheGlyAlaProCyAlaProPserSe 160  
 Db 421 AACTGGCCCTGAGACATCTGCGGAGCGGAGTTCCGGGGCGCTGCGGCCCTTCCAG 480  
 Qy 160 rCyGlyAlaThrSerGlnArgTrpSerSerGlyArpProGlnProGlnAlaProLeuTh 180  
 Db 481 TTGTGAGACCAAGCTCTCAGCGCTGAGATTCTGGAGGCCCCAGCCACAGGCCCGCTGAC 540  
 Qy 180 rSerGlySerProSerSerGlyGlyThrThrTherMetGlyTrpAlaMetProLeuMetAl 200  
 Db 541 ATCCGGCTCACCTTCTTCCAGAGGAGCCACCAAGATGGGCTGGGC-ATGCCCTTATGACC 599  
 Qy 200 aGlnGlyAlaProTrp 205  
 Db 600 CANGGCG---CCCTGG 612

RESULT 9  
 LOCUS B0694961 587 bp mRNA linear EST 15-JUL-2002  
 DEFINITION 1001104 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
 ACCESSION B0694961  
 VERSION B0694961.1 GI:21820277  
 KEYWORDS EST  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.  
 TITLE 1 (bases 1 to 587)  
 JOURNAL Yang, R.-Z., Shuldiner, A. and Gong, D.-W.  
 COMMENT EST analysis of human adipose gene expression  
 Unpublished (2002)  
 CONTACT: Gong Da-Wei  
 Division of Endocrinology, Diabetes and Nutrition  
 University of Maryland  
 660 Redwood St, HH497, Baltimore, MD 21201, USA  
 Tel: 410 706 1672  
 Fax: 410 706 1622  
 Email: dgong@medicine.umaryland.edu  
 PCR PRIMERS  
 FORWARD: CTCGGGAGCGCGCATGTTGTGCT  
 BACKWARD: AATACGACTCACTATAGGCGGAATTGG  
 Seq primer: CTTCGTACCCGGGAAATTC.  
 Location/Qualifiers

FEATURES  
 source  
 1..587  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /sex="Male and Female"  
 /tissue\_type="Adipose"  
 /clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
 /note="Vector: lambdaTriplex"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,226-68 Length: 587  
 Score: 1010.50 Matches: 192  
 Percent Similarity: 98.47% Conservative: 1  
 Best Local Similarity: 97.96% Mismatches: 3  
 Query Match: 35.66% Indels: 2  
 DB: 5 Gaps: 0

US-10-791-980-6 (1-520) x B0694961 (1-587)

Qy 211 ProaGArGGLyGlyAlaHisPheArgGlnaArgTrpSerLeuSerArgArg 230  
 Db 1 CCCGCCCGCGGAGACGACCTTCCAGCAAGATGACCGCTGCTCCCTAGCGCGCGCCG 60  
 Qy 231 GtArgAenLeuPheValaValaValaHisGluileGlyHisThrLeuGlyLeuThrHis 250

Db 61 GGGCGCAACCTGTTCGTGTCCTGGCGCAGAGATCGTCAACGCTTGCTCAACCCAC 120  
 Qy 251 SerProAlaProArgAlaLeuMetAlaProTyrTyrLeuArgLeuGlyArgAspAlaLeu 270  
 Db 121 TGCCCGCGCGCGCGCGCTCATGGCGCCCTACTACATCAAGAGAGCTGGGCGGAGCGGCTG 180  
 Qy 271 LeuSerTrpAspAspValaValaValaGlnSerLeuTyrGlyLeuProLeuGlyGlySer 290  
 Db 181 CTCAGCTGGAGACAGCTGCTGGCGCGTGCAGAGCCGTATGGGAAGCCCTTAGGGGGCTCA 240  
 Qy 291 ValaValaGlnLeuProGlyLeuLeuPheThrAspPheGlyuThrTrpAspSerTyrSer 310  
 Db 241 GTGGCGCTGCAGCTCCAGAAAGCTGTCACTGCTTGAAGACCTGGAGCTCTACAGC 300  
 Qy 311 ProGlnGlyArgArgProGluThrGlnGlyProLeuTyrCyHisSerSerPheAspAla 330  
 Db 301 CCCCAAGAGAGCGCCCTGMAAGCAGAGGCCCTAAATCTGCCACTCTTCTTCGATGCC 360  
 Qy 331 IleThrValaAspArgGlnGlnLeuTyrIlePheGlySerHisPheTrpGluVal 350  
 Db 361 ATCACTGTAGACAGGACACAGCACTGTACATTTTAAAGGAGGACATTCTGGAGAGGTG 420  
 Qy 351 AlaAlaAspGlyAsnValaSerGluProArgProLeuGlnGlnuArgTrpValGlyLeuPro 370  
 Db 421 GCAGCTGATGGCAAGCTCAGAGCCCGTCCACTGCAGAAAGATGGGTGGGCTGCC 480  
 Qy 371 ProenileGluAlaAlaValaSerLeuAsnAspGlyAspPheTyrPhePheGlyGly 390  
 Db 481 CCAC-ATTGAGGCTGCGGAGCTGATGATGATGATGATGATGATTTCTTCTTCAAAAGG 539  
 Qy 391 GtArgCyTrpArgPheArgGlyProLeuAspValaTrpGlyLeuPro 406  
 Db 540 GGTGATCTCTGAGAGGTTCG--GGGCCCAACAGAGTGGGCTCTCCA 585

RESULT 10  
 BMS60236 1065 bp mRNA linear EST 20-FEB-2002  
 LOCUS BMS60236  
 DEFINITION AGENCOURT 6564030 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5744676  
 5', mRNA sequence.  
 ACCESSION BMS60236  
 VERSION BMS60236.1 GI:18804497  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCES  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgs@bbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.llnl.gov  
 Plate: LLM12766 row: 9 column: 13  
 High quality sequence stop: 660.  
 Location/Qualifiers

FEATURES  
 source  
 1..1065  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5744676"  
 /tissue\_type="medulla"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_119"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source normal medulla from  
 anonymous male age 27. Library is oligo-dT primed and



QY 455 SerleuglnaaprtgylglyleprogluvalserglyalaLeuproargProasp 474  
 Db 335 AGTTGACAGACTGGGAGGAGCATCCCTGAGAGGTCAAGGGGCGCTCGGAGGCCCAT 276  
 QY 475 GlyserrilellephpharAsapapargtYTTTPAglLeuAaglnAlaIaleuGln 494  
 Db 275 GGCCTCATCATCTTTTTCAGAGATGACCGCTACCTGCGCCCTGACAGGCCCAACTGAG 216  
 QY 495 A1aThrThrserylArgrTrrPalathrgluleuprotTpmetGlyCy6TrrPhisAlaLen 514  
 Db 215 GCACCACTCTGGGCGCGCTGGGCCACCAAGCTGCCCTGGATGGGCTGCTGGCATGCCAAC 156  
 QY 515 SerglySerAlaLeuphe 520  
 Db 155 TCGGGAGCGCCCTGTTT 138  
 RESULT 12  
 LOCUS BX280995 499 bp mRNA linear EST 04-MAR-2003  
 DEFINITION BX280995 NIH MGC 121 Homo sapiens cDNA clone IMAGE5248347 ;  
 IMAGE:5248347, mRNA sequence.  
 ACCESSION BX280995  
 VERSION BX280995.1 GI:28612561  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 499)  
 Ebert, L., Heil, O., Hennig, S., Neubert, P., Parbach, E., Peters, M.,  
 Radelfot, U., Schneider, D. and Korn, B.  
 Human Unigeneset - RZPD3  
 Unpublished (2003)  
 TITLE Contact: Ina Rolfs  
 JOURNAL RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 COMMENT RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPDLib; I.M.A.G.E. cDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLib No.972)  
 http://www.rzpd.de/CloneCards/cgi-  
 bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heidenberg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available royalty-free from RZPD.  
 Contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13u, primer sequence: CGTGTAAACGACGCGCAGT.  
 FEATURES  
 source  
 1..499  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE5248347"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 3  
 fetal brains, female age 20 weeks, female age 24 weeks,  
 and male age 26 weeks. Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb. Insert size range  
 0.7-3.5 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 017. Note:  
 this is a NIH\_MGC Library."

Best Local Similarity: 98.80% Mismatches: 1  
 Query Match: 30.66% Indels: 2  
 DB: 5 Gaps: 0  
 US-10-791-980-6 (1-520) x BX280995 (1-499)  
 QY 119 ArglylylyArgrPheAlaIaleGlnGlyAaalySTPTrlyGlnIleLeuSerTyArg 138  
 Db 2 CGTAAGAAACCTTTGGCAAGCAAGGTAACAATGGTCAAGACACCTCTCTACCGC 61  
 QY 139 LeuValaantTrrProGlnIleIleu-ArgSerArgIlnPheGlyAlaP-OCyAlaProPr 158  
 Db 62 CTGGGAACCTGGCTGAGACATCTGCCGAGCGCGGAGTTCGGGCGCGTGCAGCGCC 121  
 QY 158 cSerSerCyaglyAlaThrSerGlnArgrTrrSerSerGlyArgrProGlnProGlnAlaPr 178  
 Db 122 TTCAGTTTGGAGACACGCTCTACGCTGGAGTTCTGGAGGCCCGACGACAGGCC 181  
 QY 178 OleuThrserylSerProSerSerlyGlyThrThrThrmetylTPAlaIaMetProle 198  
 Db 182 GCTGACATCCGCTCACCTTCTTCCAAAGGAGCCACACGATGGGCTGGCAATGCCCTTT 241  
 QY 198 uMetAlaGlnGlyAlaP-OTTPAglThrProPheLeuProArgrArgGlyGlnAlaIlePh 218  
 Db 242 GATGGCCAGAGGGGGCGCCCTGCGCACCGCTTC-CTGCCCGCGCGCGGAGAGCGCACTT 300  
 QY 218 eAspGlnaapGlnuArgrTrrSerLeuSerArgrArgGlyArgrAaIlePheValIle 238  
 Db 301 GCACCAATATGACGCTGTCTCTGAGCGCGCGCGCGGAGCACTGTTCGTGTGCT 360  
 QY 238 uAlaIleGlnIleGlyIleThrIleuGlyLeuThrIleSerProAlaProArgrAlaIleu 258  
 Db 361 GCGGACAGAGATCGGTACACGCTTGGCTCACCACTACCGCGCGCGCGCGCTCAT 420  
 QY 258 tAlaProTrrTrrlyArgrleuGlyArgrAlaIleuLeuSerTrrPapaPvalIleuAl 278  
 Db 421 GCGCGCTCTACACAGAGAGGCTGGCGCGCGCGCTGCTGAGCGAGCGAGTGTGCTGC 480  
 QY 278 aValGlnserLeuTrrgly 284  
 Db 481 CTTGCAAGCTCTGTTGAG 499  
 RESULT 13  
 LOCUS BX360790 958 bp mRNA linear EST 08-APR-2004  
 DEFINITION BX360790 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1075Y606 5-PRIME, mRNA sequence.  
 ACCESSION BX360790  
 VERSION BX360790.2 GI:46306660  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 958)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On May 5, 2003 this sequence version replaced gi:30378467.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 8754.1  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna78-CS0D1075BD03QPl&c=8754.r.  
 FEATURES  
 source  
 1..958  
 Location/Qualifiers





```

Db      258 GGGGGCGCCCTGCGCAGCCCTTC-CTGCCCCCGCGCGGAGGCGACTTCGACCAAGA 316
Qy      221 PGLuArGTpSerIuSerArArArArArArArArArArArArArArArArArArArArAr 241
Db      317 TGAGCGCTGCTCCCTGAGCGCGCGCGCGCGCAACCTGTTTGGGGCTGGCGCAAG 376
Qy      241 LuIleGlyHisThrLeuGlyLeuThrHisSerProAlaProArGAlaLeuMetAlaProT 261
Db      377 AGATGGGTACACGCTTGNCCTCACCCCACTCCCGCGCGCGCGCTCATGGGGCCCT 436
Qy      261 YTYTYLYArArLeuGlyArArArAlaLeuLeuSerTrpArArArValLeuAlaValGlnS 281
Db      437 ACTACAAAGAGGCTGAGCCCGCGCGCGCTGCTCAGCTMGAGCAGCGTGGCGGCAAG 496
Qy      281 eTLeuTYrGlyArArProLeuGlyGlySerValAlaValGlnLeuProGlyLyArLeuPheT 301
Db      497 GCTGTATGGAGAGCCCTTANGGGGCTTAGTGACCGTCCAGCTTCCAGTAAGTGTCTCA 556
Qy      301 hr 301
Db      557 CT 558

```

```

RESULT 15
BQ722257      1310 bp mRNA linear EST 16-JUL-2002
LOCUS      BQ722257
DEFINITION AGENCOURT 8291272 lupsk1 symphathetic_trunk Homo sapiens cDNA clone
IMAGE:6193589 5', mRNA sequence.
ACCESSION BQ722257
VERSION BQ722257.1 GI:21861154
SOURCE EST.
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE      1 (bases 1 to 1310)
JOURNAL      NIH-MGC http://mhc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Dr. James R. Lupski
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: L1AM13598 row: h column: 22
              High quality sequence start: 111
              High quality sequence stop: 344.
              Location/Qualifiers

```

## FEATURES

```

Source
1..1310
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6193589"
/sex="male"
/cisue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski symphathetic trunk"
/notes="Vector: pCMV-SPORE (Life Technologies); Site_1:
NotI Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTCTAGATCGCAGCGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2,166-53 Length: 1310
Score: 821.00 Matches: 154
Percent Similarity: 85.03% Conservative: 5
Best Local Similarity: 82.35% Mismatches: 23
Query Match: 28.97% Indels: 5
DB: 5 Gaps: 1

```

US-10-791-980-6 (1-520) x BQ722257 (1-1310)

```

Qy      326 SerSerPheArAlaIleThrValAspArGlnGlnGlnLeuTYrIlePheLySer 345
Db      83 TCCGCTCTCTCTTGTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 142
Qy      346 HisPheTrpGluValAlaAlaAspGlyAsnValSerGluProArGProLeuGlnGln 365
Db      143 CATTTCTGGAGAGTGGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 202
Qy      366 TrpValGlyLeuProProAsnIleGluAlaAlaValSerLeuAsnAspGlyAsPhe 385
Db      203 TGGGTGGGGCTGCCCCCAACATTGAGGCTGGCGAGTGTATTGATGATGAGATTTC 262
Qy      386 TyrPhePheLyGlyGlyArGlyArGlyTrpArPheArGlyProLyPProValTrpGlyLeu 405
Db      263 TACTTCTTCAAGAGGGGGTGAATGCTGAGAGGTTCCGGGGCCCAAGCCAGTGGGGTCTC 322
Qy      406 ProGlnLeuCyArArAlaGlyGlyLeuProArGHisProArAlaAlaLeuPhePro 425
Db      323 CCACAGCTGTGGCGGAGAGGGGCTGGCCCGCAGCTGAGCGCCCTTCTTTCCT 382
Qy      426 ProLeuArArGlyLeuIleLeuPheLyGlyAlaArGlyTYrValLeuAlaArGlyGly 445
Db      383 CCTTGGCCCGGCTATCTCTTCAAGGTGGCGCTACTAGTCTGGCCCGAGAGGGGA 442
Qy      446 LeuGlnValGluProTYrTYrProArGSerLeuGlnAspTrpGlyYIleProGluGlu 465
Db      443 CTGCGAGTGGAGCCCTACTACCCCGAGAGTGGAGAGCTGGGGAGGAGATCCCTGAGAG 502
Qy      466 ValSerGlyAlaLeu-ProArGProArAspGlySerIleIlePhePheArAspArTY 485
Db      503 GTTGGCCCGGCTGGCCCGAGAGCCGAGTGGCTCCATCATCTTCCAGAGATCACCGCTA 562
Qy      485 rTrp-----ArgLeuAspGlnAlaLyLeuGlnAlaThrTrhSerGlyArGTr 501
Db      563 CTGGTGCCTTCCACATATGTCTAACTGACAGAAACACTTCGGTGAAGCTCGGAGCAC 622
Qy      501 PALAThrgLuleuProTrp 507
Db      623 CGATCTACCTTGAATG 641

```

Search completed: June 14, 2005, 11:26:31  
 Job time : 4912 secs

**THIS PAGE BLANK (USPTO)**